Regulatory Toxicology and Pharmacology 144 (2023) 105472

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

journal homepage: www.elsevier.com/locate/yrtph

Extended one-generation reproductive toxicity study evaluating gardenia blue in Sprague Dawley rats

Melanie L. Foster [a,*], Debabrata Mahapatra [a], Robert R. Maronpot [b], Masayuki Nishino [c], Shuichi Chiba [c], Mihoko Koyanagi [c], Florence Burleson [d], Shim-mo Hayashi [e,f]

[a] Integrated Laboratory Systems LLC, An Inotiv Company, 601 Keystone Park Drive, Suite 200, Morrisville, NC, 27560, USA
[b] Maronpot Consulting LLC, 1612 Medfield Road, Raleigh, NC, 27607, USA
[c] San-Ei Gen, F.F.I., Inc., 1-1-11, Sanwa-cho, Toyonaka, Osaka, 561-8588, Japan
[d] Burleson Research Technologies, Inc., 120 First Flight Lane, Morrisville, NC, 27560, USA
[e] Tokyo University of Agriculture and Technology, Tokyo, 183-8509, Japan
[f] National Institute of Health Sciences, Kawasaki, Kanagawa, 210-9501, Japan

ARTICLE INFO

Handling Editor: Dr. Martin Van den berg

Keywords:
Gardenia blue
Food colorant
Food additive
Extended one-generation
Developmental
Reproductive
Neurotoxicity
Immunotoxicity
Safety assessment

ABSTRACT

Gardenia blue powder was administered at 0.5%, 2.5%, or 5.0% in feed to male and female Sprague Dawley rats in an Extended One-Generation Reproductive Toxicity Study (OECD Test Guideline 443). The dosed diet began 14 days before mating and was continued at the same concentration level for the entire study for all parental animals (P_0) and offspring (F_1). At weaning, offspring were allocated into one of 5 cohorts for different endpoints. P_0 and F_1 animals had blue urine, blue or black feces, and blue discolorations in gastrointestinal organs, mesenteric lymph nodes, and kidneys. This treatment-related finding was not considered adverse as there were no histopathologic correlates. There was a dose-related increase in sperm concentration in P_0 and F_1 males. There were dose-related increases in heart weights of F_1 postnatal day (PND) 21 males, male and female thyroid weights, and female TSH levels of PND 91 F_1 offspring, with no histopathological correlate. There were no consistent treatment-related adverse effects on any other parameters evaluated for general toxicity, reproductive toxicity, developmental neurotoxicity, or developmental immunotoxicity. The highest dietary concentration (5.0%) of gardenia blue powder was the no observed adverse effect level (NOAEL) for male and female rats at all life stages evaluated.

1. Introduction

Gardenia blue (GB) is a natural food colorant produced by adding beta-glucosidase to a mixture of iridoid glycosides obtained from the fruits of *Gardenia jasminoides* Ellis or *Gardenia augusta* Merrill and protein hydrolysates (Tsutsumiuchi et al., 2021). The principal coloring component is a 15,000–30,000 Dalton polymer with two genipin moieties with peptide side chains cross-linked by an alkene chain (Breslin et al., 2023). GB is currently approved for use in food in Japan, China, South Korea, and Taiwan and has been used for over 30 years in Japan (Ministry of Health, Labour, and Welfare, 2018; USDA, 2015; Do and Kwon, 2022; FDA Taiwan, 2013). It is approved for use in drinks, jams, jellies, pickled vegetables, fried nuts and seeds, bakery wares, candies, condiments, and other food categories at maximum levels of 0.2–1.0 g/kg in China (USDA, 2015).

Recently there has been a shift from synthetic colorants to natural colorants; however, few natural food colorants have been approved for use worldwide. Many natural food additives are available on the market based on the history of safe use without an acceptable daily intake (ADI) value. The ADI for GB in Korea is "not specified" (Ministry of Food and Drug Safety, 2018). The estimated daily intake of GB is 3 mg/person in Japan (Ueda et al., 2017). GB is being evaluated comprehensively in a chronic toxicity, carcinogenicity, teratology, and genetic toxicity test battery (Maronpot et al., 2023a, 2023b, 2023c; Breslin et al., 2023; Hobbs et al., 2018) in anticipation of the worldwide marketing as a food colorant to be widely used in the U.S., Europe, and other parts of the world. Previous studies have demonstrated no genetic toxicity, subchronic toxicity, or carcinogenicity hazards due to consumption of GB (Do and Kwon, 2022; Imazawa et al., 1996; Imazawa et al., 2000).

The purpose of this study was to perform an Extended One-

* Corresponding author. 601 Keystone Park Drive, Suite 200, Morrisville, NC, 27560, USA.
E-mail address: melanie.foster@inotivco.com (M.L. Foster).

https://doi.org/10.1016/j.yrtph.2023.105472
Received 1 April 2023; Received in revised form 11 August 2023; Accepted 19 August 2023
Available online 21 August 2023

Generation Reproductive Toxicity Study (OECD, 2018) evaluating effects of GB during specific life stages not covered by other toxicity assays and effects that may occur from pre- and postnatal exposure. The extensive range of assessments at stages of fertility, pregnancy, lactation for the P_0 animals, and prenatal and postnatal exposure through maturity for the F_1 animals permit greater confidence in the safety assessment for the product. Endpoints include general toxicity, reproductive indices, thyroid hormone levels, behavioral assessments, developmental immune function, and detailed histopathology of reproductive, nervous system, and immune tissues. Brain morphometry measurements were assessed in an expanded set of locations at PND 22 and PND 77 to evaluate key structures in the brain.

2. Materials and methods

2.1. Study design

Thirty male and 30 female Sprague Dawley rats were randomly allocated into each of 4 designated dose groups and administered either the carrier diet control or 0.5%, 2.5%, or 5.0% GB in feed (Supplementary Table 1). Feeding with the dosed diet began 14 days before mating and continued at the same concentration level for the entire study for all parental animals (P_0) and offspring (F_1). Mating occurred at a 1:1 male:female ratio for up to 14 days. Body weight measurements and clinical observations were performed at least weekly with daily cage-side observations.

On postnatal day (PND) 4, litters were randomly culled to 5 male and 5 female pups (F_1), as possible. Culled F_1 pups were removed from the litters and bled (pooled by litter) for serum thyroxine (T4) analysis. Evaluation of developmental markers (e.g., anogenital distance and presence of nipples/areola) was performed. Weekly litter weight measurements were recorded. On PND 21, F_1 animals were randomly allocated to one of 5 cohorts for evaluation of different endpoints.

The study was conducted in accordance with the U.S. Food and Drug Administration's (FDA) Good Laboratory Practice (GLP) Regulations 21 CFR Part 58 (FDA, 2002). This study was designed to satisfy the testing guideline 443: Extended One-Generation Reproductive Toxicity Study (OECD, 2018). The study was approved by the ILS IACUC and conducted in the AAALAC-accredited animal facilities at ILS. All procedures were in compliance with the Animal Welfare Act (7 USC 2131–2159) and Animal Welfare Regulations (9 CFR 1–4) (USDA, 2020), and animals were handled and treated according to the *Guide for the Care and Use of Laboratory Animals* (National Research Council, 2011).

2.2. Test animals

Hsd: Sprague Dawley® rats, 10–11 weeks old, were purchased from Envigo (Frederick, MD, USA) to provide 30 male and 30 female P_0 animals per dose group. After acclimation, P_0 animals were assigned to a dose group using a procedure that stratifies animals across groups using an integrated preclinical software (Provantis® [Instem, Staffordshire, UK]). Adult and post-weaning animals were identified by subcutaneously implanted transponder chips (Biomedic Data Systems, Inc., Seaford, DE, USA). Offspring were identified by toe tattoos from PND 1 through weaning. Animals were housed individually, except during mating, lactation, and for one week post-weaning when the litters were group-housed by sex. Animals were housed in polycarbonate cages with micro-isolator tops with absorbent heat-treated hardwood bedding (Northeastern Products Corp., Warrensburg, NY, USA) and had *ad libitum* access to Purina 5002 certified rodent meal diet and reverse osmosis treated tap water (City of Durham, NC, USA). The animals had polycarbonate tubes (Bio Serv, Flemington, NJ, USA) and/or Enviro-Dri® (Shepherd Specialty Papers, Watertown, TN, USA) in their cages except during mating, and received Bacon Yummies™ Certified (Bio Serv) once a week except during lactation. Suspended wire cages were used for mating. The room was controlled to 19–25 °C and 30–70% humidity, with a 12/12-h light/dark cycle.

2.3. Test article and doses

The test article was GB powder (Lot number 180910, San-Ei Gen, F.F.I., Inc., Osaka, Japan), a dark blue powder, CAS number 106441-42-3, analyzed to have 32.3% gardenia blue color polymer, 62.7% maltodextrin, 3.2% water, 1.8% other, and no detected genipin. The test article represents the actual commercial dietary component and was manufactured and characterized following the Food Safety System Certification (FSSC) 22000, as recognized by the Global Food Safety Initiative (GFSI). Purina certified 5002 meal rodent diet was mixed to produce concentrations of 0.5%, 2.5%, and 5.0% GB at RTI International (Research Triangle Park, NC, USA). This range of doses was selected to allow comparisons with other *in utero*, chronic, and carcinogenicity studies with GB. A high dose of 5% for a non-nutritional test article is recommended since a higher dietary percentage could potentially alter nutritional balance and cause compensatory increased food consumption (Maronpot et al., 2023c, EFSA Scientific Committee, 2011). Diet containing the specified dose level of GB was available *ad libitum* for the duration of the study for all animals.

2.4. Animal observations

2.4.1. P_0 generation and F_1 litters pre-weaning

Body weight measurements, clinical observations, and feed consumption measurements were performed at least weekly with daily cage-side observations. From feed consumption data, g of test article consumed/kg body weight/day was calculated. Vaginal smears were collected, starting on the day of exposure through confirmation of mating for P_0 females, and evaluated for stage of estrous cycle as described by Goldman et al. (2007). F_1 pup detailed clinical observations and body weights were recorded on PND 1, 4 (prior to culling), 7, 10, 14, and 21. On PND 1 and 4 (prior to culling), anogenital distance in all F_1 animals was determined. On PND 4, litters were randomly culled to 5 male and 5 female pups, as possible. Culled F_1 pups were removed from the litters and bled (pooled by litter) for serum thyroxine (T4) analysis. On PND 12, the presence of nipples and areolae in male F_1 pups was determined.

P_0 females were euthanized on Lactation Day (LD) 21, following allocation of pups to cohorts. P_0 males were euthanized at least 10 weeks after initial exposure. Following overnight fasting, blood and urine was collected from 10 male and 10 female rats per dose group, randomly selected. A complete gross necropsy, with organ weights and tissues collected for histopathology was performed on all parental animals. The number of uterine implantation sites was recorded. Sperm motility and concentration parameters were measured by computer-assisted sperm analysis (CASA) (Hamilton Thorne TOX IVOS system, Beverly, MA, USA). Sperm morphology was determined by microscopic examination of at least 200 spermatozoa per sample. F_1 animals not assigned to a cohort and euthanized on PND 21 were bled for individual serum T4 and thyroid stimulating hormone (TSH) concentration determination. The T4 Free assay is a chemiluminescent enzyme immunoassay (Siemens Immulite Analyzer) with a lower limit of quantitation (LLOQ) of 0.4 μg/dL and TSH was measured by radioimmunoassay with an LLOQ of 1.0 ng/mL. A complete gross necropsy was conducted on one male and one female per litter, as possible, with organ weights and tissues collected for potential histopathology. Any additional surplus animals were examined for external abnormalities. See Supplementary Table 2 for the number of animals that were analyzed for each endpoint.

2.4.2. F_1 generation general toxicity assessments (cohorts 1A and 1B)

One female and one male pup per litter were assigned to Cohort 1A and to Cohort 1B on PND 21, to yield at least 20 males and 20 females per dose group per cohort (Supplementary Table 2). Body weight measurements, clinical observations, and feed consumption measurements

were performed at least weekly with daily cage-side observations. The age of vaginal opening and preputial separation were determined, and the body weight on the day of attainment was recorded. The day of first estrus was determined and vaginal smears were prepared daily starting on PND 75 through termination to evaluate estrous cyclicity.

Cohort 1A animals were euthanized on PND 91–93. Following overnight fasting, urine was collected from 12 males and 12 females, and blood was collected from 10 males and 10 females per dose group. A complete gross necropsy, with organ weights and tissues collected for histopathology was performed on all Cohort 1A animals. Sperm motility, concentration, and morphology parameters were measured as described above for P_0 males. One half of the spleen was preserved for lymphocyte analysis by flow cytometry by Burleson Research Technologies, Inc. (Morrisville, NC, USA). Samples from all tissues harvested from the control and the high-dose adult Cohort 1A animals and all gross lesions from all dose groups were trimmed of any adherent tissue, histologically processed, embedded, sectioned at a thickness of 5 μm, and stained with hematoxylin and eosin (H&E). H&E-stained slides were evaluated by light microscopy. Ten male and 10 female rats randomly selected from each dose group from Cohort 1A had lymph nodes and bone marrow histopathologically evaluated from collected specimens. Differential ovarian follicular counts and atresia assessments were conducted on all control and high-dose females (Pedersen and Peters, 1968; Osman, 1985; Borgeest et al., 2002; Byskov, 1979; Dixon et al., 2014). Longitudinal and transverse sections of the uterine horns were evaluated. Detailed testicular histopathology was conducted on the F_1 males to identify potential treatment-related effects on testis differentiation and development and on spermatogenesis.

Cohort 1B male and female animals were euthanized on PND 105 ± 1 and PND 112 ± 1, respectively. A complete gross necropsy, with reproductive organ weights and tissues collected for potential histopathology was performed on all Cohort 1B animals. Additional samples of kidney, liver, and mesenteric lymph nodes were collected from 3 males and 3 females in the control group and 6 males and 6 females in the 5.0% dose group for non-GLP investigation by electron microscopy conducted at Charles River Laboratories (Durham, NC, USA) and for cryostat sectioning for enhanced darkfield microscopy/hyperspectral imaging at RTI International.

2.4.3. F_1 generation developmental neurotoxicity assessments (cohorts 2A and 2B)

At least 10 males and 10 females per dose group (one male or one female per litter) were assigned to Cohort 2A and to Cohort 2B on PND 21 (Supplementary Table 2). Body weight measurements, clinical observations, and feed consumption measurements were performed at least weekly with daily cage-side observations. The age of vaginal opening and preputial separation were determined, and the body weight on the day of attainment was recorded for Cohort 2A.

All Cohort 2A pups underwent auditory startle testing on PND 24 ± 1 using an automated acoustic startle system (SR-LAB, San Diego Instruments, San Diego, CA, USA). Each session consisted of 50 trials of a 120-dB, 20-msec pulse of white noise, with an 8-s inter-trial interval with background white noise at 73 dB. The mean response amplitude and time to maximum response on each block of 10 trials (5 blocks of 10 trials) were determined. Between PND 63–65, all animals assigned to Cohort 2A were evaluated in a functional observational battery (FOB) for their reactivity to auditory, visual, and proprioceptive stimuli, and grip strength. The observers were blinded to dose levels by using a counterbalanced order and temporary identification numbers for the rats. Observations were made: 1) while the rat was in the observation cage, 2) during removal of the rat from the observation cage, 3) while the rat was being held and examined for clinical observations, 4) as the animal moved freely about the open field, and 5) during manipulative tests. Motor activity was assessed in a 1-h block after the FOB for most animals. Some animals were tested on PND 67 or PND 69 due to equipment failure and repair. Motor activity was measured during 12 5-min intervals for a total of 60 min using an automated photobeam activity system (PAS, San Diego Instruments, San Diego, CA, USA). Activity was recorded as total activity (fine movements consisting of breaking the same photobeam consecutively, plus ambulations). Assignment of rats to motor activity sessions and individual enclosures was balanced across dose level and sex. Testing was conducted with red lights for room illumination and white noise at 70 ± 5 dBA.

On PND 77, Cohort 2A animals were anesthetized by isoflurane followed by whole body perfusion. Brain, spinal cord (cervical, thoracic, and lumbar), eyes, optic nerve, skeletal muscle, and sciatic and proximal tibial nerves were collected after fixation in 10% neutral buffered formalin. Samples from all tissues harvested from all groups were trimmed of any adherent tissue, histologically processed, embedded, and sectioned at a thickness of 5 μm. Tissues were processed in a random order to minimize impact of variations in histology processing. The control and high-dose slides were stained with H&E and evaluated by light microscopy. Multiple sections of the brain were prepared for evaluation: olfactory bulbs, cerebral cortex, hippocampus, basal ganglia, thalamus, hypothalamus, mid-brain (tectum, tegmentum, and cerebral peduncles), brain stem, cerebellum, and spinal cord. The brains from all animals were processed, in random order, for morphometric analysis, but only the slides from control and high-dose groups were stained. Morphometric analysis was performed on 3 levels of brain (3, 4 and 7) and at least 3 serial sections of each of the brain levels were obtained to attain the most homologous section for evaluation. A fourth serial section was created for H&E staining. Sections were stained with cresyl violet/Luxol fast blue (CV/LFB) to better highlight the neurons and myelin-rich white matter tracts as well as borders of neuroanatomic regions per Garman et al. (2016).

Neurohistopathology analysis was performed on all levels of each brain from the control and high-dose animals to include a total of 10 levels to ensure a comprehensive examination. Additionally, sections of the eyes and optic nerves, both transverse and longi tudinal sections of the cervical, thoracic, and lumbar spinal cord, and the sciatic and proximal tibial nerves were examined (Garman et al., 2016; OECD, 2018). CV/LFB stained slides for all animals in the control and 5.0% groups were randomized (between groups within a cohort) and blinded to minimize bias. Neuroanatomical location measurements for each level were adapted from Garman et al. (2016) and are depicted here in Fig. 1. Bilateral measurements for all locations (i.e., hippocampus) were preferred; however, unilateral measurements were taken for asymmetrical or artifactual sections. Cerebellar width (M8) was measured as the distance between the flocculus lobes, as the paraflocculus often gets damaged during dissection. The thickness of the cerebellar folia layers (M10 inner and M11 outer) were measured as an additional cerebellar measure.

On PND 22, Cohort 2B animals were anesthetized by isoflurane followed by whole body perfusion. Samples were collected and evaluated for neurohistopathology and morphometry as described for Cohort 2A.

2.4.4. F_1 generation developmental immunotoxicity assessments (cohorts 1A and 3)

Steady state splenic lymphocyte populations were determined by flow cytometry from Cohort 1A animals euthanized on PND 91–93, as described above. Results are presented as mean ± standard error of the mean using proportion of plot. Proportion of plot for each subpopulation indicates the number of events within the gate of interest out of the total events on each plot. Total leukocytes are gated within a FSC vs SSC size and granularity gate with a doublet discriminator. T-helper and T-cytotoxic cells are gated within the total $CD45^+CD3^+$ T cell population. T cells ($CD3^+$), B cells ($CD45RA^+$), and NK cells ($CD161a^+$) are gated with the total leukocyte ($CD45^+$) population.

At least 10 males and 10 females per dose group (one male or one female per litter) were assigned to Cohort 3 on PND 21 (Supplementary Table 2). Body weight measurements, clinical observations, and feed consumption measurements were performed at least

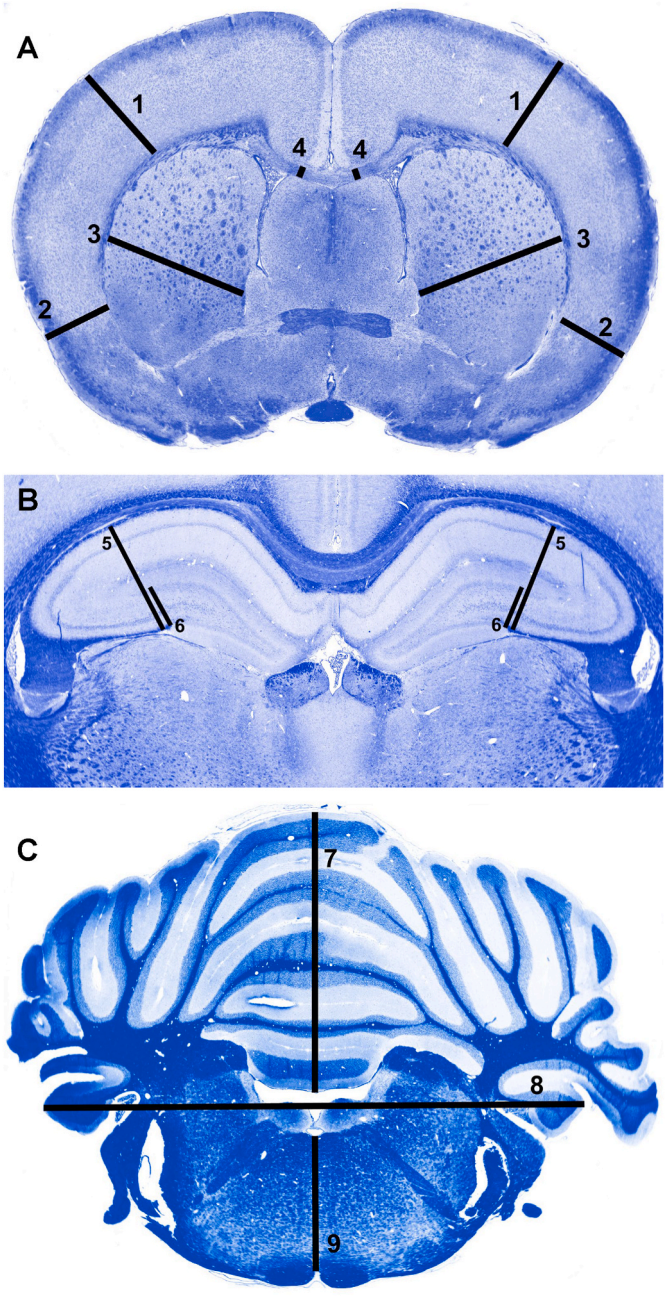

Fig. 1. A. Coronal section Level 3 showing four location measurements 1–4 (neocortex, parietal cortex, striatum and corpus callosum). Fig. 1B. Neuroanatomical level 4 showing 2 location measurements 5 & 6 (hippocampus and dentate gyrus). Fig. 1C. Neuroanatomical level 7 showing location measurements 7–9 (cerebellar height and width, and height of brain stem).

weekly with daily cage-side observations. The age of vaginal opening and preputial separation were determined, and the body weight on the day of attainment was recorded. On PND 56 ± 3 (6 days prior to euthanasia), animals were administered approximately 0.5 mL of a 4 × 10^8 Sheep Red Blood Cell suspension (SRBC; Colorado Serum Company, Denver, CO, USA, Lebrec et al., 2014) once via intravenous tail vein injection. Animals were euthanized 6 days later and blood was collected via cardiac puncture. Serum collected from Cohort 3 animals was evaluated for T cell-dependent antibody response (TDAR) using a rat anti-sheep red blood cell IgM ELISA (Life Diagnostics, Inc., West Chester, PA, USA) at Burleson Research Technologies, Inc. following GLPs.

2.5. Statistical analysis

Group means and standard deviations were calculated and reported. Statistical analyses for all endpoints (final body weight, body weight gain, food consumption [g/kg/day], absolute and relative tissue weights, neurotoxicological endpoints, developmental endpoints, reproductive indices, urinalysis endpoints, clinical pathology endpoints, and histopathological endpoints [if possible]) were conducted using Statistical Analysis System version 9.2 (SAS Institute, Cary, NC, USA). Clinical pathology endpoints with reported values below the lower limit of quantitation were assigned a value of 50% of the lower limit, and those with values above the upper limit of quantitation were assigned a value of 150% of the upper limit for statistical analysis. Anogenital distance was normalized to the cube root of body weight for individual pups. Where appropriate, statistical analysis evaluated litter effects. The significance threshold for all statistical tests was $\alpha = 0.05$.

For continuous endpoints with data for multiple treatment groups measured post-weaning, data were first evaluated for possible outliers by reviewing studentized residual plots. Data were then evaluated for equality of variances using Levene's test. If the Levene's test p-value was significant, the data were transformed and re-evaluated for equality of variances. Data were transformed and re-evaluated sequentially, first with log, then square root, and finally inverse, and the first transformation that achieved equality of variances was used for subsequent analyses. Raw or transformed data with equivalent variances were analyzed with a one-way ANOVA, Dunnett's test for multiple comparisons and one-way ANOVA contrasts for linear trend. Data that did not have equivalent variances, even after transformation, were analyzed with Dunn's test for multiple comparisons and the Jonckheere-Terpstra trend test.

For continuous endpoints with data for multiple treatment groups measured pre-weaning, data were first evaluated for possible outliers and equality of variances as described above. Raw or transformed data with equivalent variances were analyzed with an Analysis of Covariance (ANCOVA), with Dam ID as a covariate to control for litter effects, Dunnett's test for multiple comparisons and ANCOVA contrasts for linear trend. If an endpoint did not have equality of variances between groups, subsequent analyses were conducted on litter means normalized by litter size. Normalized litter means were then analyzed following the post-weaning statistical workflow for continuous endpoints with data for multiple treatment groups.

For continuous endpoints with data for only one treatment group, data were compared to the appropriate concurrent vehicle control group by first evaluating the data for normality with the Shapiro-Wilk test. If the Shapiro-Wilk p-value was significant, the data were transformed and re-evaluated sequentially, first with log, then square root, and finally inverse; and the first transformation that achieved normality was used for subsequent analyses. Raw or transformed data that were normally distributed were analyzed with a *t*-test. Data that were not normally distributed, even after transformation, were analyzed with the Mann-Whitney test (also known as a Wilcoxon rank-sum test).

Endpoints in the FOB using graded or count scales were analyzed using a non-parametric strategy. When 75% or fewer of the scores for an endpoint in all groups were tied, the non-parametric Kruskal-Wallis Test was used to analyze the data, and in the event of a significant result ($p \leq 0.05$), Dunn's Test was used to compare the treated groups with the control group. When more than 75% of the scores in any group were tied, Fisher's Exact Test was used to compare the proportion of ties in the groups. Endpoints in the FOB using descriptive or quantal scales were analyzed using Fisher's Exact Test, regardless of data distribution.

Data from the motor activity test and acoustic startle test, with repeated measurements within a session, were analyzed with both sexes using a repeated measures ANOVA. A significant effect ($p \leq 0.05$) in that test can appear as an effect of dose (a difference between groups in the total across all measurements in a session) or as an interaction between dose and time (a difference between groups at specific measurement

periods). If the dose effect was significant, pairwise comparisons of treatment groups with the appropriate control group were conducted using Dunnett's Test. If the dose × time interaction was significant, pairwise comparisons of treatment groups with the appropriate control using Dunnett's Test were performed at each measurement period within the repeated measures model.

3. Results

3.1. Gardenia blue intake

All dose formulations met the acceptance criteria for concentration and homogeneity. There were no significant differences in food consumption for P_0 males. Mean food consumption in P_0 females exposed to 5.0% GB was significantly increased compared to the control group during Gestation Days 0–21 (a significant linear dose trend was also noted). There were no significant differences in food consumption for P_0 females during the pre-mating or lactation periods. There was a significant trend of increased food consumption in Cohort 1A F_1 males, but no other dose-related differences in food consumption for F_1 animals (Supplementary Table 3).

Food consumption on a per kg body weight basis was higher in females than males, especially during lactation for P_0 females, and in younger F_1 animals. This corresponded to increased test article consumption (Supplementary Table 4). Mean test article consumption was 226.6, 1150.2, and 2349.7 mg/kg/body weight/day for P_0 males in the 0.5%, 2.5%, and 5.0% GB exposure groups, respectively. Test article consumption during the pre-mating, gestation, and lactation periods for P_0 females was similar to that of males and increased by dose, although higher, reflecting the higher food consumption per kg body weight in females, especially during lactation. Mean test article consumption was 354.6, 1730.0, and 3420.7 mg/kg body weight/day during pre-mating; 335.7, 1683.0, and 3461.2 mg/kg body weight/day during gestation; and 989.1, 4981.2, and 9855.9 mg gardenia blue/kg body weight/day during lactation for P_0 females in the 0.5%, 2.5%, and 5.0% GB exposure groups, respectively. Mean test article consumption was 3914.6 and 4349.3 mg/kg body weight/day for male and female Cohort 1A F_1 animals in the 5.0% GB group. There were no treatment-related differences in food efficiency for P_0 or F_1 animals, except for a linear trend of increased efficiency for F_1 males in Cohort 2A.

3.2. P_0 generation

No adverse clinical observations, morbidity, or mortality were observed in the P_0 animals. There were no effects on reproductive success for P_0 males or females (Supplementary Table 5). A statistically significant increasing trend was noted for P_0 total sperm concentration (millions/gram), with no significant pairwise differences. There were no changes in body weight for the P_0 males (Fig. 2A) and there were dose-related trends of decreased body weight at the end of lactation for the P_0 females (Fig. 2B). Blue-colored urine and blue or black feces were present in the animals fed GB. Urinary protein in females had a significant increasing trend, with the 5.0% group being significantly increased compared to the control group. Male and female urine bilirubin and female urobilinogen (Supplementary Table 6) had significant increasing dose trends and significant increases for the 2.5% and 5.0% female groups compared to the control group. There were no other indications of hepatic damage or anemia to correlate with these findings. Highly colored red urine can produce false positives for bilirubin by interfering with the colorimetric assay (Hoilat and John, 2022). With the lack of corroborating results in the blood or by histopathology, it seems likely that the urine bilirubin and urobilinogen elevations are in fact influenced by the blue-colored urine. Some of the changes in urine values in females (i.e., urinary protein) may be attributed to the incidences of spontaneous chronic progressive nephropathy.

There were no dose-related changes in hematology. Statistically significant clinical chemistry changes in P_0 animals included a decreasing linear trend for creatinine in males (with a significant pairwise difference between the 2.5% GB and control groups), increasing linear trend for glucose in males (no significant pairwise differences), decreased alanine aminotransferase (ALT) in males exposed to 2.5% GB (no significant dose-response trend), an increasing linear trend for gamma glutamyl transferase in females (no significant pairwise differences), and an increasing linear trend for TSH in males (with a significant pairwise difference between the 5.0% GB and control groups) (Supplementary Tables 7 and 18-20). These levels were within our historical control data range, except that the male controls and 5.0% ALT values were higher than our historical control data. None of these findings were correlated with any histopathological changes.

There were statistically significant changes in the absolute and/or relative (organ to body weight ratio) weights and increasing or

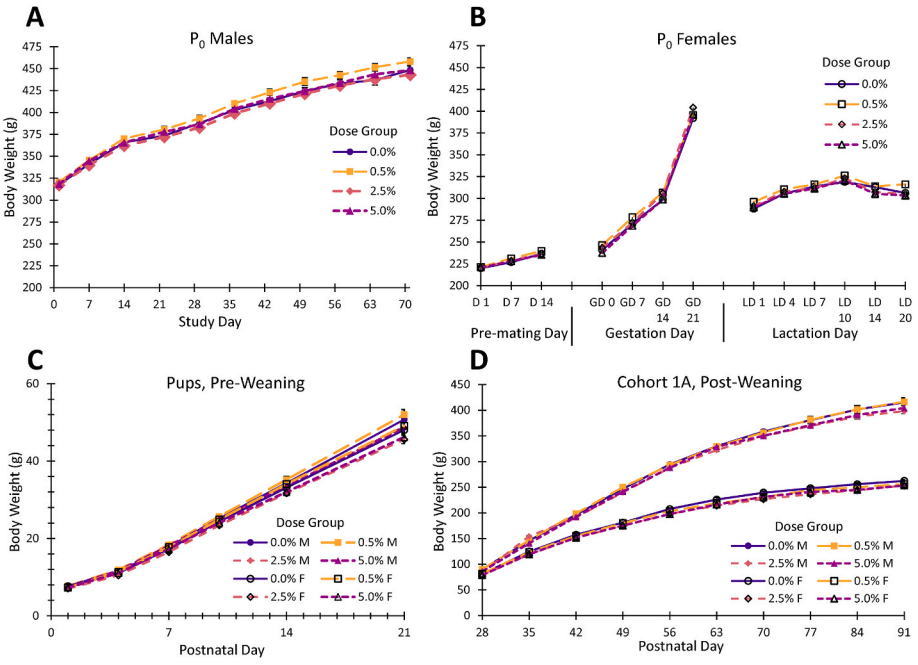

Fig. 2. P_0 and F_1 Body Weights (mean ± SEM). Fig. 2A. P_0 Male Body Weights. There were no differences in P_0 male body weights. Fig. 2B. P_0 Female Body Weights. There were dose-related trends of decreased body weight at the end of lactation. Fig. 2C. Pre-weaning body weights in F_1 animals (Litter averages by sex). There were trends of dose-related reductions in body weights at PND 21 for both sexes. Fig. 2D. Cohort 1A Post-weaning Body Weights. There were trends of decreased body weights at several days, with no pairwise differences from control.

decreasing trends in several organs in P_0 males and females (Supplementary Tables 21-24). Decreasing trends in male absolute and relative spleen weights were noted, along with a significantly decreased mean relative weight in males exposed to 5.0% GB compared to control. Thyroid gland weights were decreased in males (decreasing linear trend in absolute weights, with significantly lower mean thyroid weight at 2.5% compared to control) but were significantly increased in females at 5.0% compared to the control group, with increasing linear trends in absolute and relative weights. Mean relative adrenal weight in males exposed to 0.5% GB was significantly lower than the control (no overall dose-response trend), but in females there was a significant increasing trend in relative adrenal weight (with no significant pairwise differences). A significant increasing trend was also noted for relative kidney weight in females (with no significant pairwise differences). There were no correlating histopathological changes observed in these organs. The only gross pathology finding considered to be related to test article administration was blue discoloration of internal organs (i.e., mesenteric lymph nodes, duodenum, jejunum, ileum, cecum, colon, and stomach) and blue or dark discoloration of the kidney. The blue or dark discoloration of the internal organs observed grossly was not apparent microscopically. No dose-related histopathology was present.

3.3. F_1 generation

There were no dose-related differences in survival of F_1 female offspring during the pre-weaning period (PND 0, 4 [pre-cull], and 21) (Supplementary Table 8). There was a significant increase ($p = 0.02$) in the number of dead male offspring from PND 1–4 in the 5.0% GB group, which were nearly the same as for females, but the control male group had zero deaths during that period. The litter mean and standard deviation for the number of dead pups from PND 1–4 was 0.3 ± 0.6, 0.3 ± 0.7, 0.2 ± 0.4, 0.3 ± 0.5 for females; and 0.0 ± 0.0, 0.1 ± 0.4, 0.0 ± 0.2, 0.3 ± 0.6 for males, in the 0.0%, 0.5%, 2.5%, and 5.0% GB groups, respectively. The average numbers of pup deaths were within the range of our historical control data. There were significant dose-related trends of reductions in pup body weights at PND 21 and weight gains from PND 1–21 for both sexes during the pre-weaning period, although there were no pairwise differences in comparison to the control group (Fig. 2C). F_1 males and females that were exposed to GB at test concentrations of 0.5% or greater had blue to black feces generally throughout the treatment period. These changes in feces color were attributed to the dark blue color of the test article and were not adverse. No other test article-related clinical findings were noted.

There were no significant differences in mean measured or normalized anogenital distance on PND 1 and 4 for male or female pups or the number of male pups with retained nipples on PND 12 in the GB-exposed groups when compared to the control group (Supplementary Table 9). Preputial separation and vaginal opening were not affected, other than a decrease in body weight at vaginal opening for 2.5% females, without a dose-response trend. There were no significant differences in T4 levels on PND 4 for pups or in T4 and TSH levels on PND 21 for surplus male and female pups (Supplementary Table 10). Statistically significant organ weight changes in surplus PND 21 male pups included a decreasing linear trend for heart weight in males with a significant pairwise difference between the 2.5% and 5.0% GB groups when compared to the control group (Supplementary Tables 25 and 26). Significant decreasing trends were noted in organ weights for kidneys, brain, epididymides, and testes in surplus PND 21 males, but there were no significant pairwise differences between GB-exposed groups and the control group. There were no significant differences in relative organ weights for either sex or absolute organ weights for female PND 21 pups when compared to the control group. These changes in absolute organ weights may reflect the decreased body weight found at PND 21 in the 5.0% group.

3.4. F_1 generation general reproductive/developmental toxicity assessments (cohorts 1A and 1B)

Statistically significant decreasing trends in mean body weight were noted for F_1 Cohort 1A males on PND 28 and 35 and Cohort 1A females on PND 35, 42, 49, 56, and 84, with no significant pairwise differences noted at any of these timepoints between the test article-exposed groups and the respective control group (Fig. 2D). There were no other statistically significant differences in weekly body weights and weight changes for F_1 males and females from weaning through termination. The number of days from vaginal opening until first estrus were unaffected by GB exposure (Supplementary Table 9). There was a dose-related trend of increased regular estrous cycles in Cohort 1A only, without any group differences from the control group (Supplementary Table 10). The mean estrous cycle length was significantly reduced in the 0.5% group of Cohort 1B only, when compared to control group animals in the same cohort, without a dose-responsive trend. The mean cycle length was within normal ranges. Increased regular cyclicity and a mean cycle length of 4.6 days are not adverse effects. Statistically significant higher F_1 Cohort 1A mean total sperm concentrations (millions/gram and millions/mL) were noted for the 5.0% GB group males compared to the control group, with a corresponding increasing dose-response trend for the millions/mL parameter. Statistically significant urinalysis changes in F_1 Cohort 1A animals were limited to decreasing linear trends for urine volume and urinary ketones and an increasing trend for bilirubin in males (but no pairwise differences were noted). There were no significant urinalysis changes in Cohort 1A females. Statistically significant hematology changes in F_1 Cohort 1A animals were limited to a decreasing linear trend in absolute neutrophils for males and an increasing linear trend in platelets for females. No significant pairwise differences were noted between the GB-exposed groups and each respective control group (Supplementary Tables 27 and 28). These changes were considered minor and not indicative of an adverse test article-related effect. Statistically significant clinical chemistry changes in F_1 Cohort 1A animals consisted of a decreasing linear trend in albumin for males (with no significant pairwise differences) and an increasing linear trend in TSH in females, with a significantly increased mean value in the 5.0% GB-exposed females compared to the control group (Supplementary Table 10). These values were all within our historical control data range.

There were statistically significant changes in the absolute and/or relative (organ to body weight ratio) weights and increasing or decreasing trends in several organs in F_1 Cohort 1A males and females (Supplementary Table 11, remaining in Supplementary Tables 30–33). Thyroid gland weights (absolute and relative) were significantly increased in males and females following exposure to 0.5%, 2.5%, and 5.0% GB compared to control values, with corresponding significant increasing linear dose trends for male absolute and relative weights and female relative weights. There was not a significant correlation of thyroid weight with increased TSH levels in females and no change in T4 levels in males and females. Significant increasing linear trends in pituitary gland weights (absolute and relative), with significantly higher mean absolute and relative pituitary weights only at 0.5% compared to controls, were noted for males and females. A significant decreasing trend in absolute mesenteric lymph node weights was noted for females (no significant pairwise differences were observed). Significantly higher mean relative testes weight was noted for the 2.5% GB-exposed males compared to the control group (no significant linear trend was observed). There were no correlating histopathological changes observed in these organs. Statistical analysis revealed a significant increase in the numbers of medium-sized ovarian follicles in the 5.0% dose group (Supplementary Table 12). This increase was considered incidental since corresponding numbers for primordial follicles and/or large follicles were not significantly elevated. Also, no associated changes were observed in the ovarian weights, clinical observation parameters or ovarian histology to support the increased numbers of medium-sized

follicles in this group. Hence, this change was not considered test article related. Blue discoloration of internal organs was noted as a gross observation, but not microscopically, similar to the P_0 animals. No treatment-related histopathological findings were present. There were no dose-related changes in organ weights in Cohort 1B animals. Evaluation of samples of kidney, liver and mesenteric lymph nodes from a subset of control and 5.0% male and female animals by electron microscopy revealed no differences and no abnormalities.

3.5. F_1 generation developmental neurotoxicity assessments (cohorts 2A and 2B)

All Cohort 2A pups underwent auditory startle testing on PND 24 ± 1 (Supplementary Tables 34 and 35). There was a significant effect ($p < 0.0001$) of block for the maximum response amplitude, which indicates that habituation to the repeated stimulus occurred. The group by sex, group by block, and group by block by sex interactions were not significant for the maximum response amplitude or the time to maximal response.

Between PND 63–65, all Cohort 2A animals were evaluated in an FOB for their reactivity to auditory, visual, and proprioceptive stimuli and grip strength (Supplementary Tables 36 and 37). The only statistically significant effects found in the FOB were an increased forelimb grip strength in male rats in the 5.0% GB group with a dose-related trend, an increased number of rears in females in the 0.5% GB group, and a trend of increased urine in females, without any pairwise differences compared to control. The grip strength values and the number of rears are within our historical control values. Grip strength measurements are not a direct measure of muscle strength and require motivation by the animal to grasp the grid. Grip may be affected by drugs affecting motivation, emotional state or muscle relaxation (Gauvin, 2021). No other indications of these effects were observed.

The total movements of male and female control rats approached an asymptote during the last 10–15 min of the 60-min motor activity testing session (Fig. 3), as specified in the criteria for the test guidelines. The interval number was a significant factor ($p < 0.0001$) indicating habituation had occurred, as expected and required. There was no significant dose group by sex effect, so analyses were done with both sexes combined. There were no significant effects on total movements as a repeated measure, or on the session cumulative total activity analyzed with a two-way ANOVA.

No statistically significant changes in F_1 Cohort 2A (PND 77) brain or body weights were observed in any animal in this cohort (Supplementary Table 13). The only significant change in morphometric brain measurements was an increase in thickness of the outer layer of cerebellar folia in males. (Supplementary Table 14). All protocol-required tissues from control and high-dose animals in F_1 Cohort 2A were microscopically evaluated. This included neurohistopathological evaluation of 10 coronal sections (1 through 10) of the brain; 3 transverse and longitudinal sections each from the cervical, thoracic, and lumbar spinal cord; peripheral nerves (sciatic and proximal tibial); and eyes with optic nerves. There were no microscopic findings considered to be test article-related. Histologic change (one incidence of mineralization) observed in the periocular tissue was considered incidental.

There were statistically significant decreases in the total body weight and corresponding increases in the relative brain to body ratio for F_1 Cohort 2B females in the mid-dose (2.5%) and high-dose (5.0%) groups (Supplementary Table 13). However, no statistically significant changes were observed for the post-fixed brain weights in these animals. In addition, there was a statistically significant increase in the cerebellar height in the females at 5.0% GB compared to controls of this cohort. There were no other statistically significant changes in morpho metric brain measurements in the F_1 Cohort 2B (PND 22) animals (Supplementary Table 15).

3.6. F_1 generation developmental immunotoxicity assessments (cohort 3)

There were no statistically significant differences in splenic $CD8^+$ T cell or $CD4^+$ T cell proportions in GB-exposed rats from Cohort 1A at any dose level compared to rats fed a control diet (0.0% GB) (Supplementary Table 16). Total leukocyte ($CD45^+$), T cell ($CD3^+$), B cell ($CD45RA^+$), and NK cell ($CD161a^+$) proportions were also not altered by any dose level of GB.

In Cohort 3 animals that were administered SRBC suspensions on PND 56, female anti-SRBC IgM antibody production appeared to decrease with increasing GB concentration; however, the data was variable and no statistically significant differences between treatment groups were observed (Supplementary Table 17). There were also no statistically significant differences in anti-SRBC IgM antibody concentrations between treatment groups in males only or with both sexes combined.

4. Discussion and conclusions

Overall, there were minimal treatment-related effects in the parental and offspring generations. There were no adverse clinical observations, morbidity, or mortality in the parental animals. Reproductive success for parental males or females was unaffected. There was a dose-related increase in sperm concentration and no effects on the estrous cycle. Organ weight changes were sporadic, with a decrease in male thyroid and spleen weights and an increase in TSH, and an increase in female thyroid weights, with no change in TSH. In an embryo-fetal developmental toxicity study with oral gavage dosing of GB in Sprague Dawley rats (Breslin et al., 2023), there were no changes in maternal T3, T4, or TSH,

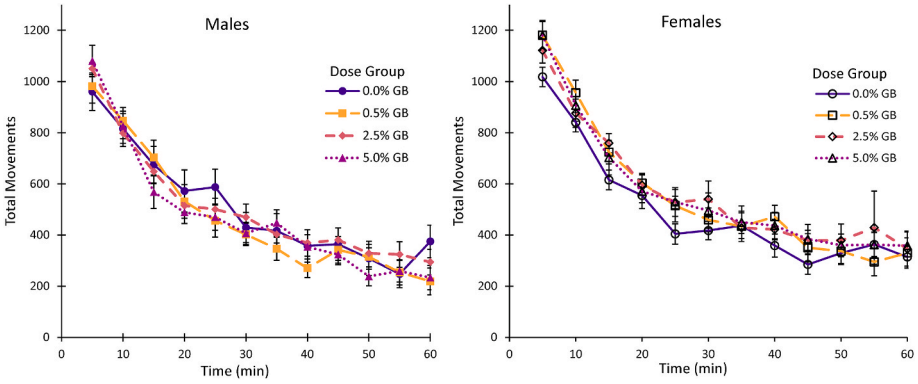

Fig. 3. Summary Motor Activity movements per 5-min interval (Mean ± SEM) by Sprague Dawley rats administered Gardenia Blue. Fig. 3A. F_1 Cohort 2A Males Fig. 3B. F_1 Cohort 2A Females There were no significant effects on total movements as a repeated measure, or on the session cumulative total activity analyzed with a two-way ANOVA.

or thyroid weights. There were no consistent changes in clinical pathology measures, other than an increase in urine bilirubin, urobilinogen, and blue-colored urine in P_0 females exposed to 2.5% and 5.0% GB. The blue urine may cause a false positive result for bilirubin and urobilinogen in urine. Gross observations of blue gastrointestinal tract and mesenteric lymph nodes, with blue or dark kidneys, are considered treatment-related but not adverse and do not correlate with any histopathological observations. Test article consumption during the pre-mating and gestation periods for P_0 females was similar to that of males and unchanged by dose, although slightly higher, reflecting the higher food consumption per kg body weight in females. The NOAEL for P_0 males is 5.0% GB (2349.7 mg/kg/body weight/day) and 5.0% for P_0 females (3420.7, 3461.2, or 9855.9 mg/kg body weight/day during pre-mating, gestation, and lactation).

In the F_1 offspring generation, there was a small statistically significant increase in the number of male pup deaths between PND 1 and 4. This difference is likely due to the fact that there were no deaths of male pups in the control group during this time period. There were no treatment-related deaths in the post-weaning period. There was a trend of dose-related decrease in body weight and bodyweight gain at weaning, but no group differences. Dose-dependent trends of decreased bodyweights at some timepoints only in Cohort 1A were not observed in the other cohorts, indicating that this was a spurious finding. There were no dose-responsive changes in developmental landmarks. Similar to P_0 males, sperm concentration was increased at 5.0% GB exposure. Estrous cycle effects were inconsistent and not adverse. Gross observations of blue gastrointestinal tract and mesenteric lymph nodes, with blue or dark kidneys, similar to those observed in the parental generation are considered treatment-related but not adverse and do not correlate with any histopathological observations. Heart weights were decreased in the 2.5% and 5.0% GB males, with no corresponding clinical pathology or histopathology. Increases in pituitary weights were present in males and females administered 0.5% GB. Increased thyroid weights at all treatment levels and a dose-responsive trend of increased TSH at 5.0% GB in PND 91 females do not correlate with any other clinical chemistry or histopathology. In considering potential effects on the hypothalamic pituitary axis, multiple regression analysis was conducted and there was no correlation of TSH levels with thyroid, pituitary, brain, or liver weights. There were no effects on T4 or TSH at younger ages or in the P_0 females and no corresponding reproductive, behavioral, clinical chemistry, or histopathology changes. In a similar *in utero* exposure study with GB in Sprague Dawley rats (Maronpot et al., 2023c), there were no effects on TSH, T3, or T4 in approximately 90-day-old offspring and inconsistent thyroid weight effects in 52-week-old offspring. Taken together, the TSH and thyroid weight changes in PND 91 females do not represent a clear or consistent toxicological response. No histopathological findings were treatment-related. At PND 22, terminal body weight was decreased and relative brain weight was increased in females administered 2.5% and 5.0% GB. Cerebellum height was also increased at 5.0%. These effects did not persist to PND 77, but the thickness of the outer layer of cerebellar folia was increased in males at 5% GB. There were no indications of treatment-related brain pathology. Cohort 2A animals tested for auditory startle at PND 24 $\pm$ 1 had no group differences and testing for motor activity and the FOB at PND 64 $\pm$ 5 revealed no differences in motor activity and a few differences in the FOB. There were no indications of immunotoxicity, by steady state splenic lymphocyte populations or by TDAR.

Conduct of an Extended One-Generation Reproductive Toxicity Study evaluated potential effects of GB during specific life stages not covered by other toxicity assays and effects that may occur from pre- and postnatal exposure. The NOAEL for F_1 males and F_1 females is 5.0% GB powder in feed (3914.6 mg/kg bodyweight/day for Cohort 1A males and 4349.3 mg/kg bodyweight/day for Cohort 1A females), based on a lack of dose-responsive adverse effects. This level is the same as determined in previous subchronic and carcinogenicity studies in adult F344 rats (Imazawa et al., 1996, 2000). The NOAEL for GB consumption in this study for P_0 and F_1 animals represents approximately 50,000 to 200,000-fold estimated daily consumption for adults in Japan (Ueda et al., 2017). Teratology studies conducted in rats and rabbits showed no evidence of teratogenic effects at the limit dose of 2000 mg/kg/day GB (Breslin et al., 2023). The NOAEL was also 5.0% in feed for adult Sprague Dawley rats administered GB in feed for 12 or 24 months in a combined chronic toxicity and carcinogenicity study (Maronpot et al., 2023b), and in a carcinogenicity study with rasH2 mice (Maronpot et al., 2023a). These study results are also supported by an *in utero* assessment of Sprague Dawley rats in which animals were administered GB in feed as parental animals and offspring, with a NOAEL of 5.0% (Maronpot et al., 2023c). Completion of this comprehensive test battery assures confidence in the safety assessment for gardenia blue for people at all stages of life.

Funding body information

Gardenia Blue Interest Group.
1-1-11, Sanwa-cho, Toyonaka.
Osaka, 561–8588, Japan.

Final decisions on the collection, analysis, and interpretation of data, the writing of the report, report content, and in the decision to submit the article for publication remained with the performing lab (Melanie Foster and Debabrata Mahapatra) and the study monitor (Robert Maronpot).

CRediT authorship contribution statement

Melanie L. Foster: Investigation, Data curation, Formal analysis, Supervision, Writing – original draft, Writing – review & editing. **Debabrata Mahapatra:** Investigation, Writing – review & editing. **Robert R. Maronpot:** Investigation, Writing – review & editing. **Masayuki Nishino:** Conceptualization, Project administration, Writing – review & editing. **Shuichi Chiba:** Conceptualization, Project administration, Writing – review & editing. **Mihoko Koyanagi:** Conceptualization, Project administration, Writing – review & editing. **Florence Burleson:** Investigation. **Shim-mo Hayashi:** Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests which may be considered as potential competing interests: M.F. and D.M. are employees of Integrated Laboratory Systems, LLC, an Inotiv company (ILS). The Gardenia Blue Interest Group funded the conduct of this study at ILS. M.N., S.C., and M.K. are employees of San-Ei Gen F.F. I., Inc., Japan. This company manufactures food-related products containing colorants. R.M. is a consultant for San-Ei Gen F.F.I., Inc., Japan.

Data availability

Summary data have been included as Supplementary data tables.

Acknowledgements

This study was supported by funding from the Ministry of Agriculture, Forestry and Fisheries of Japan to the Gardenia Blue Interest Group under the project for promotion of export environment for “the project for preparation of an application of gardenia blue among those for existing additives”.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.yrtph.2023.105472.

References

Borgeest, D., Symonds, D., Mayer, L.P., Hoyer, P.B., Flaws, J.A., 2002. Methoxychlor may cause ovarian follicular atresia and proliferation of the ovarian epithelium in the mouse. Toxicol. Sci. 68, 473–478. https://doi.org/10.1093/toxsci/68.2.473.

Breslin, W.J., Mesnard, J.L., Maronpot, R.R., Koyanagi, M., Chiba, S., Nishino, M., Hayashi, S., 2023. Prenatal developmental toxicity of gardenia blue, a natural food colorant, in rats and rabbits. Toxicol. Res. Appl. 7, 1–14. https://doi.org/10.1177/23978473231169090.

Byskov, A.G., 1979. Atresia. In: Midgley, A.R., Sadler, W.A. (Eds.), Ovarian and Follicular Development and Function. Raven Press, New York, pp. 41–57.

Dixon, D., Alison, R., Bach, U., Colman, K., Foley, G.L., Harleman, J.H., Haworth, R., Herbert, R., Heuser, A., Long, G., Mirsky, M., Regan, K., Van Esch, E., Westwood, F. R., Vidal, J., Yoshida, M., 2014. The international harmonization of nomenclature and diagnostic criteria for lesions in rats and mice. Nonproliferative and proliferative lesions of the rat and mouse female reproductive system. J. Toxicol. Pathol. 27, 1S–107S. https://doi.org/10.1293/tox.27.1S.

Do, B., Kwon, H., 2022. Genotoxicity test of eight natural color additives in the Korean market. Gene Environ. 44 (1), 19. https://doi.org/10.1186/s41021-022-00247-0.

EFSA (European Food Safety Authority) Scientific Committee, 2011. Guidance on conducting repeated-dose 90-day oral toxicity in rodents on whole food/feed. EFSA J. 9, 2438.

Food and Drug Administration (FDA Taiwan), 2013. Sanitation Standard for Natural Edible Colorants. https://consumer.fda.gov.tw/Law/Detail.aspx?nodeID=518&lang=1&lawid=139.

Garman, R.H., Li, A.A., Kaufmann, W., Auer, R.N., Bolon, B., 2016. Recommended methods for brain processing and quantitative analysis in rodent developmental neurotoxicity studies. J. Toxicol. Pathol. 44 (1), 14–42. https://doi.org/10.1177/0192623315596858.

Gauvin, D.V., 2021. The functional observational battery: utility in safety assessment of new molecular entities. In: Llorens, J., Barenys, M. (Eds.), Experimental Neurotoxicology Methods, Neuromethods, 172. Humana Press, New York, pp. 165–198. https://doi.org/10.1007/978-1-0716-1637-6_8.

Goldman, J.M., Murr, A.S., Cooper, R.L., 2007. The rodent estrous cycle: characterization of vaginal cytology and its utility in toxicological studies. Birth Defects Res B 80, 84–97. https://doi.org/10.1002/bdrb.20106.

Hobbs, C.A., Koyonagi, M., Swartz, C., Davis, J., Maronpot, R., Recio, L., Hayashi, S., 2018. Genotoxicity evaluation of the naturally-derived food colorant, gardenia blue, and its precursor, genipin. Food Chem. Toxicol. 118, 695–708. https://doi.org/10.1016/j.fct.2018.06.001.

Hoilat, G., John, S., 2022. Bilirubinuria. National Library of Medicine, National Center for Biotechnology Information. NBK557439. StatPearls Publishing, Florida. https://www.ncbi.nlm.nih.gov/books/NBK557439/.

Imazawa, T., Nishikawa, A., Furukawa, F., Tanakamaru, Z., Lee, I.S., Kim, H.C., Takahashi, M., 1996. A 13-week subchronic toxicity study of gardenia blue in F344 rats. Eisei Shikenjo Hokoku 114, 27–32 (Japanese). PMID: 9037861.

Imazawa, T., Nishikawa, A., Furukawa, F., Kasahara, K., Ikeda, T., Takahashi, M., Hirose, M., 2000. Lack of carcinogenicity of gardenia blue colour given chronically in the diet to F344 rats. Food Chem. Toxicol. 38 (4), 313–318. https://doi.org/10.1016/S0278-6915(99)00166-0.

Lebrec, H., Molinier, B., Boverhof, D., Collinge, M., Freebern, W., Henson, K., Mytych, D., Ocks, H., Wange, R., Yang, Y., Zhou, L., Arrington, J., Christin-Piche, M., Shenot, J., 2014. The T-cell-dependent antibody response assay in nonclinical studies of pharmaceuticals and chemicals: study design, data analysis, interpretation. Regul. Toxicol. Pharmacol. 69, 7–21. https://doi.org/10.1016/j.yrtph.2014.02.008.

Maronpot, R., Koyanagi, M., Hayashi, S., Mahapatra, D., Nishino, M., Iniwa, M., 2023a. Gardenia blue is not carcinogenic in the rasH2 mouse. Toxicol. Res. Appl. 7 https://doi.org/10.1177/23978473231173093.

Maronpot, R., Ramol, Y., Nyska, A., Sproul, C., Moore, R., Koyanagi, M., Chiba, S., Nishino, M., Hayashi, S., 2023b. Chronic toxicity and carcinogenicity study of dietary gardenia blue in Sprague Dawley rats. Food Chem. Toxicol. 176, 113734, 10.106/j.ct.2023.113734.

Maronpot, R., Streicker, M., Mahapatra, D., Moore, R., Koyanagi, M., Chiba, S., Nishino, M., Hayashi, S., 2023c. Twelve-month *in utero* safety assessment of gardenia blue, a natural food colorant. J. Toxicol. Pathol. 36, 171–179. https://doi.org/10.1293/tox.2023-0030.

Ministry of Food and Drug Safety: Korea, 2018. Study on the Re-evaluation of Health Based Guidance Values for Food Additives. https://doi.org/10.1186/s41021-022-00247-0.

National Research Council, 2011. Guide for the Care and Use of Laboratory Animals, eighth ed. The National Academies Press, Washington, DC. https://doi.org/10.17226/12910.

OECD (Organisation for Economic Co-operation and Development), 2018. *OECD Guidelines For the Testing Of Chemicals, Section 4*, Test No. 443: Extended One-Generation Reproductive Toxicity Study. OECD Publishing, Paris. https://doi.org/10.1787/20745788.

Osman, P., 1985. Rate and course of atresia during follicular development in the adult cyclic rat. J. Reproduction Infertil. 73, 261–270. https://doi.org/10.1530/jrf.0.0730261.

Pedersen, T., Peters, H., 1968. Proposal for a classification of oocytes and follicles in the mouse ovary. J. Reproduction Infertil. 17, 555–557. https://doi.org/10.1530/jrf.0.0170555.

The Ministry of Health, Labour and Welfare, The Consumer Affairs Agency, 2018. Japan's Specifications and Standards for Food Additives (JSFA), ninth ed. http://www.nihs.go.jp/dfa/dfa-j/shokuten_kikaku_j.html.

Tsutsumiuchi, K., Toyoshima, T., Hasegawa, F., Terasawa, R., Honda, W., Sakakibara, M., Ishida, Y., Ikai, Y., Ishibashi, R., Furuya, K., Morimoto, T., Ishizuki, K., Nishizaki, Y., Masumoto, N., Sugimoto, N., Sato, K., Oka, H., 2021. Molecular structure of gardenia blue pigments by reaction of genipin with benzylamine and amino acids. J. Agric. Food Chem. 69 (13), 3904–3911. https://doi.org/10.1021/acs.jafc.0c07948.

U. S. Department of Agriculture (USDA), 2020. Animal Welfare Act and Animal Welfare Regulations (APHIS 41-35-076). 7 USC 2131-2159 (Animal Welfare Act) and 9 CFR 1-4 (Animal Welfare Regulations). Retrieved from. https://www.aphis.usda.gov/animal_welfare/downloads/AC_BlueBook_AWA_508_comp_version.pdf.

Ueda, Y., Seguro, K., Okano, H., Sato, K., 2017. Study on Estimation of Food Additives Intake Based on Statistical Survey of Production Volume (The 6th Survey on Existing Food Additives)., 37. National Institute of Health Sciences, JAFAN, pp. 1–38.

U.S. Food and Drug Administration (FDA), 2002. Good Laboratory Practice for Nonclinical Laboratory Studies, 21 CFR 58 (43 FR 60013, 22 December 1978, last amended in 67 FR 35729, 21 May 2002). Retrieved from. https://www.ecfr.gov/current/title-21/chapter-I/subchapter-A/part-58.

USDA Foreign Agricultural Service, 2015. Global Agricultural Information Network, Chinese Standards for Food Additives - GB2760-2015. https://apps.fas.usda.gov/newgainapi/api/report/downloadreportbyfilename?filename=Standard%20for%20Food%20Additive%20Use%20-%20GB2760-2015_Beijing_China%20-%20Peoples%20Republic%20of_4-28-2015.pdf.