



Reversal Learning Performance in the XY* Mouse Model of Klinefelter and Turner Syndromes

Shawn M. Aarde^{1*}, Haley Hrcir¹, Arthur P. Arnold¹ and James D. Jentsch²

¹ Department of Integrative Biology and Physiology, Laboratory of Neuroendocrinology of the Brain Research Institute, University of California, Los Angeles, Los Angeles, CA, United States, ² Department of Psychology, Binghamton University, Binghamton, NY, United States

OPEN ACCESS

Edited by:

Carsten T. Wotjak,
Max Planck Institute of Psychiatry
(MPI), Germany

Reviewed by:

Francesco Papaleo,
Istituto Italiano di Tecnologia, Italy
William Davies,
Cardiff University, United Kingdom

*Correspondence:

Shawn M. Aarde
saarde@binghamton.edu

Specialty section:

This article was submitted to
Pathological Conditions,
a section of the journal
Frontiers in Behavioral Neuroscience

Received: 02 July 2019

Accepted: 19 August 2019

Published: 06 September 2019

Citation:

Aarde SM, Hrcir H, Arnold AP
and Jentsch JD (2019) Reversal
Learning Performance in the XY*
Mouse Model of Klinefelter and Turner
Syndromes.
Front. Behav. Neurosci. 13:201.
doi: 10.3389/fnbeh.2019.00201

Klinefelter syndrome (KS; 47, XXY) and Turner syndrome (TS; 45, XO) are caused by two relatively common sex chromosome aneuploidies. These conditions are associated with an increased odds of neuropsychiatric disorders, including attention deficit/hyperactivity disorder (ADHD), as well as impairments in cognition that include learning delays, attentional dysfunction and impulsivity. We studied cognitive functions in the XY* mouse model, which allows comparison of XXY to XY males (KS model), and XO to XX females (TS model). We evaluated adult mice with and without gonads, using a version of an operant reversal-learning task (RLT) that can be used to measure various facets of learning, impulsivity and attention. In the KS model, only one measure related to impulsivity – perseverative responding under reversal conditions – reliably discriminated gonadally intact XXY and XY mice. In contrast, a fundamental learning impairment (more trials to criterion in acquisition phase) in XXY mice, as compared to XY, was observed in gonadectomized subjects. No other task measures showed differences consistent with KS. In the TS mouse model, XO mice did not show a pattern of results consistent with TS, similar to past observations. Thus, the application of this RLT to these XY* models reveals only limited behavioral impairments relevant to KS.

Keywords: reversal learning, Klinefelter syndrome, Turner syndrome, animal model, XY*, sex chromosome aneuploidy, perseveration

INTRODUCTION

Sex-chromosome aneuploidies, such as Klinefelter syndrome (KS; 47, XXY) and Turner syndrome (TS; 45, X), are associated with various cognitive impairments and with an increased risk of psychiatric and/or neurodevelopmental disorders, including attention-deficit hyperactivity disorder, autism, schizophrenia, and bipolar disorder (Russell et al., 2006; Cederlöf et al., 2014; Belling et al., 2017; Tartaglia et al., 2017; Zhao and Gong, 2017). For example, an increased risk of ADHD has been observed in people with either KS (Cederlöf et al., 2014) or TS (Russell et al., 2006). Moreover, KS men who do not have a diagnosis of autism spectrum disorder nonetheless have higher Autism-spectrum Quotient scores (AQ) (Baron-Cohen et al., 2001), including poorer scores on the “attention switching” subscale, indicating subclinical neurobehavioral manifestations that will require quantitative, rather than diagnostic, behavioral approaches (van Rijn et al., 2012b, 2014; Cederlöf et al., 2014).

Cognitive assessments indicate a modestly lower full-scale IQ (5–10 points) in both KS and TS, with higher performance IQ relative to verbal IQ in KS, and the opposite relationship in TS (Hong and Reiss, 2014). Moreover, in a recent assessment that included a Stroop Word-Color Test and the

Wisconsin Card Sorting Test, KS men exhibited an greater Stroop effect (thus, poorer inhibition) and made more errors (including more perseverative errors) on the card sort task relative to controls (thus, poorer behavioral flexibility) (Skakkebaek et al., 2013).

To better understand the mechanisms by which sex-chromosome aneuploidies impact phenotypes, various animal models of KS or TS have been developed (Lue et al., 2010; Wistuba, 2010; Cox et al., 2014; Burgoyne and Arnold, 2016). Behavioral studies of these models indicate that both KS and TS aneuploidies impact learning, impulsivity and/or attention (Isles et al., 2004; Davies et al., 2005, 2007; Lue et al., 2005; Lewejohann et al., 2009; Chen et al., 2013b; Cox et al., 2015). In order to better resolve the validity and usefulness of such models, we chose to study an operant spatial reversal-learning task (RLT) using the XY* mouse model on a C57BL/6J background, which produces both KS-like and TS-like genotypes and their controls in the same litters.

In the XY* mouse model, the Y chromosome has a modified pseudoautosomal region that recombines abnormally with the X chromosome. This generates four male gametes: X, Y*, X^{Y*} (a single chromosome produced by the fusion of X and Y*), and Y*^X (which paradoxically is not a Y chromosome, but is analogous to an X chromosome massively deleted for most X genes, leaving predominantly the minute pseudoautosomal region) (Eicher et al., 1991; Burgoyne and Arnold, 2016). Thus, mating an XY* male to an XX female produces mice that are the near-equivalent of XX, XY (i.e., XY*), XXY (i.e., 40,XX^{Y*}), and XO (i.e., 40,XY*^X) (Eicher et al., 1991; Burgoyne et al., 1998; Wistuba et al., 2010; Burgoyne and Arnold, 2016). Although 39, XO mice have been generated and studied (Isles et al., 2004; Davies et al., 2005, 2007; Lopes et al., 2010; Raznahan et al., 2013), the current study used 40,XO mice (i.e., 40,XY*^X) on a C57BL/6J background because the 39,XO model is not available on that background.

Although the data are limited, mouse models of KS do appear to have some face validity in terms of learning impairments (Rovet et al., 1996; Ross et al., 2008; Skakkebaek et al., 2017). For example, in a novel-object recognition test, XXY mice from the XY* model failed to show the typical pattern of more exploration of the novel object relative to the more familiar object, but control XY mice did (Lewejohann et al., 2009). In a different KS model, 41,XXY mice were slower to acquire Pavlovian appetitive approach behavior than XY mice (Lue et al., 2005).

Regarding TS mouse models, performance on a serial RLT (Y-maze-based, visual, non-spatial) was worse in 39,XO than 40,XX mice, but only when the X-chromosome in the 39,XO mice was of maternal origin (Davies et al., 2005). Thus, 39,XO mice possessing the maternal X have face validity for the general learning impairments observed in TS (Garron, 1977; Loesch et al., 2005; Mazzocco, 2006). However, as the relevance of imprinting to learning impairments in TS remains unclear (Lepage et al., 2012), so does its relevance to the validity of mouse models of TS (Davies et al., 2005, 2007).

On the 5-choice serial reaction time task (5-CSRTT; tests visual attention), 39,XO mice displayed poorer accuracy and

slower reaction times compared to 40,XX mice (outbred MF1 strain background) (Davies et al., 2007). However, the presence of the Y*^X chromosome in 40,XO mice rescued the attention deficit and there were no parent-of-origin effects detected. Additionally, on measures of impulsivity, 39,XO and 40,XX mice did not differ. Thus, the 39,XO mouse appears to have face validity in terms of the attentional deficits observed in TS (McCauley et al., 1987; Deloos et al., 1993; Ross et al., 2002; Russell et al., 2006; Green et al., 2018; Mauger et al., 2018), but not for the impulsivity observed in TS (Romans et al., 1997; Tamm et al., 2003; Russell et al., 2006).

To further study behavior and cognition in the XY* model, we compared XXY males to XY males (KS comparison), as well as XO females (40,XO; i.e., XY*^X) to XX females (TS comparison), on an operant RLT (Laughlin et al., 2011; Linden et al., 2018). This task was chosen because it measures some of the constructs impaired in KS and TS (learning, impulsivity, and attention), as well as for its relevance to the current literature.

A recent advance in the behavioral analysis of this particular version of the reversal learning task (Laughlin et al., 2011; Linden et al., 2018) makes use of a *a priori* defined and empirically determined “change points” (CP) in the learning curve to better resolve within-phase behavior dynamics (Gallistel et al., 2004; Papachristos and Gallistel, 2006; Linden et al., 2018). In feedback-based operant-learning tasks, the acquisition curve of an individual subject is typically not one of a negatively accelerating, gradual progress toward a performance asymptote (Gallistel et al., 2004; Papachristos and Gallistel, 2006). Rather, such learning curves are typically an artifact of group averaging and thus fail to represent the behavioral dynamics of individuals which include changes in the rate of learning caused by prolonged intervals of stable performance and/or abrupt, large step-like changes in proficiency – i.e., “change points” (Gallistel et al., 2004; Papachristos and Gallistel, 2006).

However, a recent report made successful use of these CP to analyze data from a spatial RLT in rats (Klanker et al., 2015). Specifically, they split trials into those occurring before (PRE) or after (POST) the largest empirically determined CP in the individual learning curves. They observed that, for rats that learned the reversal, phasic post-reward dopamine release in the ventromedial striatum was lower POST than PRE change-point. Additionally, post-cue dopamine release was higher on trials that followed a rewarded trial than on those prior rewarded trials – but, only PRE change-point. Thus, we made use of this analytical tool as it appears to better resolve discrete phases in learning.

Some effects of KS and TS are potentially attributed to altered levels of gonadal hormones. KS men have lower levels of testosterone than XY men (Gravholt et al., 2018; Klein et al., 2018), and TS women have altered ovarian function (Ross et al., 1995, 2004; Romans et al., 1998; Rovet, 2004; Klein et al., 2018). Accordingly, we analyzed the role of gonadal hormones on task performance by comparing gonadectomized mice to gonad-intact controls.

Under the assumptions that the XXY and XO groups of the XY* model are sufficiently valid models of KS and TS, respectively, we predicted a pattern of genotypic effects on behavior similar to that seen in people with KS and

TS, with the exception of premature responding in the XO vs. XX comparison for which no difference was predicted based upon prior performance on the 5-CSRTT/1-CSRTT (Davies et al., 2007). We expected that if genotypes differed on behavioral measures, that some of the differences might disappear if they depended on group differences in levels of gonadal hormones. However, sex-hormone replacement therapies in adulthood do not appear to remediate the cognitive deficits observed in people with KS (Fales et al., 2003; Kompus et al., 2011; Liberato et al., 2017; Skakkebaek et al., 2017) or TS (Ross et al., 2002, 2004; Klein et al., 2018), and the cognitive consequences of reducing/blocking sex hormones in people with either KS or TS, or in this animal model, are unknown.

MATERIALS AND METHODS

Subjects

XY* (XY-Star) Mouse Model

Subjects were adult (100–129 days old at the time of testing) mice from the XY* model (Eicher et al., 1991; Burgoyne et al., 1998; Burgoyne and Arnold, 2016), backcrossed from Jackson lab strain 2021 to C57BL/6J for at least 13 generations. Offspring were produced by mating XY* males to XX females. This model produces four genotypes – XY*X, XX, XY* and XX^{Y*} – that are the near-equivalent of XO, XX, XY, and XXY, respectively (Eicher et al., 1991; Burgoyne et al., 1998; Chen et al., 2013a; Cox et al., 2014; Burgoyne and Arnold, 2016) (**Supplementary Figure 5B**). The Y*X chromosome is similar to an X chromosome with massive deletion of most genes, leaving the pseudoautosomal region and a few nearby genes (Burgoyne and Arnold, 2016), so that XY*X females have X monosomy (40, XO) except for the presence of a second PAR. In this model, XY and XXY mice have testes, and XO and XX mice have ovaries. In XXY (40, XXY) mice, one X chromosome and the Y chromosome are fused end-to-end (Burgoyne and Arnold, 2016).

Gonadectomy (performed under 2–5% isoflurane anesthesia and carprofen analgesia using aseptic procedures; gonads were exposed, clamped, ligated, then excised), or a control sham surgery, were performed at 72–99 days of age (mean = 82), and behavioral testing began 25–41 (mean = 30) days later. The eight groups were: XY,GDX ($N = 12$); XY,SHAM ($N = 11$); XXY,GDX ($N = 15$); XXY,SHAM ($N = 16$); XX,GDX ($N = 13$); XX,SHAM ($N = 14$); XO,GDX ($N = 14$); XO,SHAM ($N = 14$).

Genotyping

The number of X chromosomes and the presence of the Y chromosome were determined by fluorescent *in situ* hybridization that labeled both the X chromosome oncogene Rab9b and the entire Y chromosome in interphase lymphocytes using the Kreatech KI-30505 kit (Leica Biosystems). Genotype was verified by phenotypic measurements of anogenital distance and, for the two groups of males, by visual assessment of testis size; males with one X chromosome have testes roughly six times larger than males with two X chromosomes (Wistuba, 2010).

Husbandry

After gonadectomy, mice were transferred to a vivarium that varied in temperature (69–79°F) and humidity (~20–~70%), under a 14 h:10 h light:dark cycle. Behavioral testing was performed 1–4 h before onset of the dark phase. Mice were housed in groups of 2–4 mice/cage, on sawdust bedding.

To facilitate motivation to perform the task, all mice were food-restricted to maintain a body weight of 82.5% (range of 80–85%) of baseline weights (i.e., the weight just prior to the start of food restriction). Careful control of body weight was achieved by measuring body weight at the same time each day (4 h before dark onset) and by adjusting their daily allotment of standard rodent chow in response to all of the following: (1) overnight changes in body weight, (2) the difference between the current and target weight (82.5%), and (3) the amount of food consumed during training/testing. To prevent mice from dropping below 80%, the magnitude of chow adjustment was asymmetric with 5x-larger adjustments when below than when above 82.5%. Additionally, when a single ration shared among cagemates did not suffice to maintain target weights, mice were briefly separated (1–3 h) by a cage divider before providing appropriate allocations. Chow was provided to subjects 30 min after testing. Water was always available except during testing (~1 h), and the post-testing interval before being given their chow. All experimental procedures were approved by the UCLA Chancellor's Animal Research Committee.

Operant Conditioning Testing – Reversal-Learning Task

Overview, Equipment

Operant chambers (Med Associates; St. Albans, VT, United States) were housed within sound-attenuating cubicles and were equipped with 5 horizontally arranged nose-poke apertures on one curved wall, a reinforcer-delivery magazine on the opposite wall, a white-noise generator (~85 dB; always on), and a house light (located outside and above the magazine-side of the chamber). Nose-poke apertures and the magazine were illuminated via recessed lamps, and entries into apertures/magazine were detected via interruption of an infrared-beam sensor. Data were collected via MedPC IV software (Med Associates) running custom behavioral programs. The progression of reversal-learning task phases and the range of days on each phase are illustrated in **Supplementary Figure 5A**.

Habituation Phase

After initiation of food restriction, mice were transferred to the operant testing room where the equipment was turned on. Mice were placed singly into clean home cages (no bedding, two nestlets). After ~30 min, a ceramic ramekin containing 1.4 g of the reward (14-mg Dustless Precision purified reinforcer pellets; BioServ item #F05684) was placed into the cage. After 1 h, the mice were returned to their original home cages. This was repeated for a total of 4 consecutive days.

Magazine Training

Mice were next given two sessions of magazine training, during which reinforcer pellets were delivered on a modified

variable-time schedule ($30 \pm 0, 5, \text{ or } 10 \text{ s}$) in which reinforcers were delivered every $\sim 30 \text{ s}$, so long as a nose-poke into the magazine was detected before the time of the next pellet delivery (i.e., cessation of magazine poking suspended the timer and thus stopped further pellet deliveries until poking resumed). The magazine was illuminated when the pellet was delivered and then darkened when an entry into the magazine was detected. Sessions ended after delivery of 60 pellets or after 60 min, whichever occurred first.

Observing-Response Training

Mice were next trained to produce a sustained duration nose-poke response into the central aperture (hole 3 of 5 within the horizontal array); this was termed an “observing” response, for use in the acquisition and reversal phases of testing. If the mice were successful at sustaining their response for the required duration, a pellet was delivered into the magazine, the magazine light was illuminated, and the center aperture was darkened. Failure to sustain the response caused a 2-s timeout during which the house light was illuminated and the center aperture was darkened. The required holding duration (RHD) was block randomized and possible durations were incremented in the next session if the mouse met performance criterion as follows (failure to meet criterion resulted in a repeat of the condition):

- (a) RHD = 1, 5, or 10 csec, criterion ≥ 30 pellets in $< 60 \text{ min}$.
- (b) RHD = 10, 20, or 30 csec, criterion ≥ 40 pellets in $< 60 \text{ min}$.
- (c) RHD = 20, 30, or 40 csec, criterion ≥ 40 pellets in $< 60 \text{ min}$.
- (d) RHD = 30, 40, or 50 csec, criterion ≥ 50 pellets in $< 60 \text{ min}$.

Acquisition Learning Phase (ACQ)

Mice were next trained in $\sim 1\text{-h}$ daily sessions to acquire a simple spatial discrimination. Each trial began with illumination of hole 3 of 5; upon completion of the observing response (RHD = 1, 10, or 20 csec), the center aperture was darkened and the flanking apertures (holes 2 and 4 of 5) were illuminated. For each mouse, one of the two holes was *a priori* assigned as the correct side (designated correct aperture was counterbalanced across groups), and responses into that aperture would lead to pellet delivery into an illuminated magazine. Reward retrieval would darken the magazine and trigger the start of a 3-s inter-trial interval (ITI) after which the center aperture was re-illuminated and the next trial commenced.

Responses into the designated incorrect aperture triggered a 5-s timeout, during which all aperture lights were extinguished and the house light was illuminated. Alternatively, failure to nose-poke into either of the lit target apertures within 30-s of their illumination was counted as a response omission and led to a 5-s timeout. Timeouts for incorrect responses and omissions were followed by a 3-s ITI. Timeouts for observing-response failures continued as scheduled in observing-response training. However, after these timeouts, the ACQ trial continued until a successful observing-response triggered target presentation.

Premature responses (responses to the target apertures *before* target presentation) and extraneous observing responses (EOR – repetitive responses into the central aperture *during* target presentation) were recorded but had no scheduled consequences.

Subjects continued daily testing until they reached a performance criterion of 80% correct within a sliding window of 20 trials within a single session.

Reversal Learning Phase (REV)

The reversal phase began the day following successful completion of the acquisition phase. In the reversal phase, the apertures designated as correct and incorrect were swapped, but all other contingencies remained the same as those in the acquisition phase – including the performance criterion. Testing continued until criterion was met.

Behavioral Measures and Dependent Variables

Learning Rate

The total number of trials required to reach performance criteria, and the pattern of correct and incorrect choices made over the progression to criteria, were evaluated to characterize individual learning behavior as a function of testing phase. Additionally, to characterize within-phase performance changes, each phase was bisected on the trial in which the maximum change in learning occurred as determined by calculating the maximum linear deviation of the cumulative correct response curve (Gallistel et al., 2004; Klanker et al., 2015), leading to a PRE and POST change-point fraction of trials within each testing phase (illustrated in figure 5 of Klanker et al., 2015). Calculations were performed as described in Gallistel et al. (2004) (illustrated in figure 5 of Gallistel et al., 2004). Briefly, for each trial, a linear deviation was calculated (i.e., the difference between the y -value of the *curve* of cumulative correct responses and the y -value of the straight *line* connecting the first and last values of that curve). Then, the trial on which this valued maximized was designated the 1st POST trial.

Choice Accuracy by Prior Choice (Behavioral Flexibility Score)

Choice behavior was evaluated as a function of feedback from the preceding trial (reward vs. timeout punishment) to quantify the tradeoff between behavioral flexibility (changing response after punishment; “Lose-Shift”) and stability (repeating response after reward; “Win-Stay”). Thus, we calculated the proportion of “Lose” and “Win” trials that were correct [$p(\text{Lose-Shift})$ and $p(\text{Win-Stay})$, respectively], then we calculated a flexibility-stability difference score [FS, where $\text{FS} = p(\text{Lose-Shift}) - p(\text{Win-Stay})$]. Thus, a positive FS would indicate better accuracy from behavioral flexibility than stability. Moreover, baseline differences in overall error rates due to other factors are subtracted out by this tradeoff calculation.

Error Parameters (Perseverative and Regressive Errors)

In addition to error rate (errors per trial), two measures parameterized the shape of the learning curve which, in the reversal phase, was characterized by an initial period of perseveration on the old rule followed by a period characterized by occasional “regressive” errors. For the 1st measure, we calculated the maximum number of consecutive incorrect

responses (**MAXCI**) as an operational definition of perseverative responding. For the 2nd, we calculated the proportion of the learning curve marked by *regression* from – rather than progression toward – the performance criterion. Learning curves often show variability in these transient periods of decreased proficiency (Gallistel et al., 2004). This **Regress Score** was calculated by identifying blocks of 5 trials in which performance accuracy declined relative to the preceding block, summing the values associated with these declines in accuracy $[-\sum_{i=5}^n (\% \text{ correct next 5 trials} - \% \text{ correct prior 5 trials}) i]$, then normalizing this sum to the total number of trials.

Premature Responses

Premature responses (nose pokes into the unlit target apertures outside of the period of target presentation) and the number of timeouts/trial incurred from observing response fails (i.e., failures to sustain the observing response until the target stimuli are presented) were parsed as a function of testing **Phase** (acquisition vs. reversal; ACQ vs. REV), and the within-phase interval relative to the maximum CP in the learning curve (PRE vs. POST). Premature responses were also analyzed with respect to response **Side** (correct side in acquisition or correct side in reversal; C_{ACQ} or C_{REV}).

Response Latencies

Target-response latency was the time from target onset until the start of the 1st nose-poke into an illuminated target aperture. Reward-retrieval latency was the time from pellet delivery to the start of 1st subsequent nose-poke into the magazine. Trial-initiation latency was the time from the end of inter-trial interval to the 1st observing response (regardless of whether it was long enough to trigger target presentation).

All three latency measures were analyzed as a function of Testing **Phase** (acquisition vs. reversal; ACQ vs. REV), and the within-phase interval relative to the maximum CP in the learning curve (PRE vs. POST). Target-Response latencies were parsed by response **Side** (correct side in acquisition or correct side in reversal; C_{ACQ} or C_{REV}). Trial-initiation latencies were additionally analyzed by the outcome of the **Prior Trial** (rewarded or unrewarded).

Extraneous Observing Responses

During target-stimuli presentation, additional nose-pokes into the center aperture were recorded but were without scheduled consequences. We analyzed EOR as a function of testing phase and change point.

Omissions

Target stimuli were presented for 30 s. Failure to respond to one of the target apertures within this interval caused a timeout and an omission to be recorded. We analyzed omissions per trial as a function of testing phase and change point.

Statistical Design

Because one of our central hypotheses focused on gonadal males with one vs. two X chromosomes, we directly compared the

performance of gonadally intact XY vs. XXY mice. Separately, we compared XO vs. XX mice as a secondary test of the effects of X chromosome number in the presence of ovaries. Measures were analyzed under designs that included the between-subjects factors of X-chromosome number (i.e., **X-dose**; **1X** vs. **2X**) and gonadectomy group (**GDX**; **GDX** vs. **SHAM**), as well as the within-subjects factors of testing **Phase** (acquisition vs. reversal; **ACQ** vs. **REV**), **Prior-Trial Outcome** (rewarded vs. unrewarded), and before/after the maximum change in the learning curve [i.e., CP; **PRE** vs. **POST**; determined by calculating the maximum change in the cumulative correct response curve (Gallistel et al., 2004; Klanker et al., 2015)]. Premature responses and target-response latencies were analyzed with reference to response **Side** (correct side in ACQ or correct side in REV; C_{ACQ} or C_{REV}).

Data Analyses

Measures were analyzed by the Generalized Estimating Equations (GEE) procedure available in IBM SPSS Statistics version 25 using the robust estimator of the unstructured covariance matrix, maximum likelihood parameter estimation, Type III model effects, and the Chi-square Wald score for full log quasi-likelihood function. A normal distribution with identity link function was used except when the assumptions of normality and linearity were violated (normality determined by Kolmogorov–Smirnov Test (KST), linearity determined by examination of P-P plots). In those cases, a standard data transformation (link function in GEE; options included untransformed, log, or square root) was chosen based on the lowest Chi Square/highest *p*-value of KSTs of transformed data. Two measures – premature responses/trial, and omissions/trial – could not be normalized by these standard transformations. However, they both adequately fitted, and were thus analyzed using a gamma distribution with log-link. Significant model effects ($p < 0.05$) comprising more than two means were delineated with *post hoc* paired-means comparisons under a Bonferroni correction. Figures plot the raw (untransformed) data, with the exception of premature responses, which were plotted as the value of $\log_{10}(y_i + 1)$.

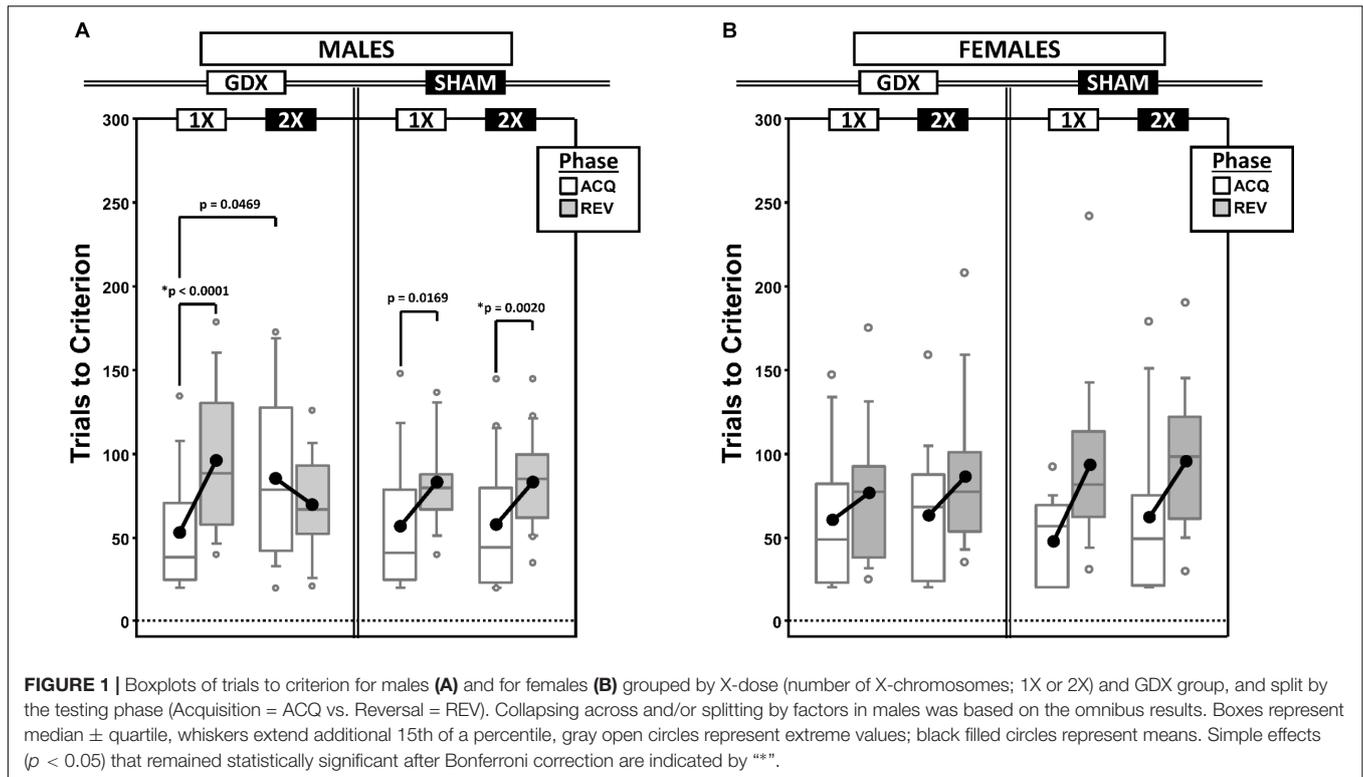
RESULTS

Trials to Criterion

As a measure of learning rate, we counted the total number of response trials required to reach the preset performance criterion (omissions were not counted). Before counting, trials were first parsed by testing phase (acquisition vs. reversal: ACQ vs. REV) and by a derived intra-phase CP in the learning curve (PRE vs. POST) (**Figures 1A,B** and **Supplementary Figures 1A,B**).

Males

Generalized estimating equations analysis confirmed significant effects of Phase [$\chi^2_{(1)} = 15.935$, $p < 0.0001$] and CP [$\chi^2_{(1)} = 43.899$, $p < 0.00001$], but there was no Phase*CP interaction [$\chi^2_{(1)} = 0.000$, $p = 0.997$]. As expected, more trials



were required in REV than ACQ and there were more trials PRE than POST (**Supplementary Figure 1A**).

There were two significant interactions involving experimental groups (**Figure 1A**): X-dose*Phase [$\chi^2_{(1)} = 7.938$, $p = 0.0048$] and X-dose*GDX*Phase [$\chi^2_{(1)} = 7.856$, $p = 0.0051$]. *Post hoc* comparisons confirmed a significant simple effect of Phase in both XY-GDX mice ($p < 0.0001$) and XXY-SHAM mice ($p = 0.0020$), but the simple effect of Phase in both XY-SHAM mice ($p = 0.0169$) and XXY-GDX mice ($p = 0.18$), as well as the simple effect of X-dose (XXY > XY) in GDX mice in ACQ ($p = 0.0469$) did not survive Bonferroni correction ($p > 0.06$ for all other simple effects). Thus, although XY-GDX, XY-SHAM, and XXY-SHAM groups showed the expected reversal cost on trials to criterion, the XXY-GDX mice required *fewer* trials in REV than ACQ due to both relatively more trials in ACQ and fewer trials in REV as compared to the other groups, likely reflecting impaired acquisition of the initial rule in this group.

Females

Generalized estimating equation analysis confirmed effects of Phase [$\chi^2_{(1)} = 21.064$, $p < 0.00001$] and CP [$\chi^2_{(1)} = 16.384$, $p < 0.0001$], as well as a Phase*CP interaction [$\chi^2_{(1)} = 4.188$, $p = 0.0407$]. *Post hoc* comparisons confirmed that the effect of Phase was significant both PRE ($p < 0.00001$) and POST ($p = 0.005$). However, the effect of CP was significant for REV ($p = 0.0001$), but not for ACQ ($p = 0.152$). As expected, there were more trials PRE than POST, and more in REV than ACQ (**Supplementary Figure 1B**). GEE analysis did not confirm any significant effects or interactions involving X-dose or GDX (all $p > 0.1$) (**Figure 1B**).

Error Rate

Errors per trial were parsed by testing phase (acquisition vs. reversal; ACQ vs. REV) and the intra-phase CP in the learning curve (PRE vs. POST), the latter to determine within-phase dynamics of learning (**Supplementary Figures 1C,D**).

Males

Generalized estimating equation analysis confirmed significant effects of Phase [$\chi^2_{(1)} = 45.172$, $p < 0.000001$] and CP [$\chi^2_{(1)} = 551.751$, $p < 0.000001$], as well as a Phase*CP interaction [$\chi^2_{(1)} = 18.932$, $p = 0.00001$]. As expected, there were more errors made PRE than POST, regardless of phase, and more errors in REV than ACQ, regardless of CP (all $p < 0.001$; **Supplementary Figure 1C**). There were no significant effects or interactions involving X-dose or GDX (all $p > 0.1$).

Females

Generalized estimating equation analysis confirmed effects of Phase [$\chi^2_{(1)} = 23.670$, $p < 0.00001$] and CP [$\chi^2_{(1)} = 610.346$, $p < 0.00001$], and a Phase*CP interaction [$\chi^2_{(1)} = 6.873$, $p = 0.0088$]. As expected, there were more errors made PRE than POST, regardless of phase, and more errors in REV than ACQ, regardless of CP (all $p < 0.0001$; **Supplementary Figure 1D**). There were no significant effects or interactions involving X-dose or GDX (all $p > 0.1$).

Behavioral Flexibility Score (FS)

To score the tradeoff between behavioral flexibility and stability, we calculated the difference between the proportion of responses that were correct following punished trials [p(Lose-Shift); index

of flexibility] vs. those that were correct following rewarded trials [$p(\text{Win-Stay})$; index of stability] [$\text{FS} = p(\text{Lose-Shift}) - p(\text{Win-Stay})$; **Supplementary Figures 2A,B**]. Thus, a positive FS score would indicate better accuracy from behavioral flexibility than stability. These calculations were based on observations parsed by both the testing phase (ACQ vs. REV) and the CP in the learning curve (PRE vs. POST; used to examine within-phase dynamics).

Males

As expected, mean FS decreased from a positive value PRE to a negative value POST in both ACQ and REV (**Supplementary Figure 2A**), reflecting a shift from a Lose-Shift strategy to a Win-Stay strategy. GEE analysis confirmed significant main effects of CP [$\chi^2_{(1)} = 38.214$, $p < 0.00001$] and Phase [$\chi^2_{(1)} = 9.825$, $p = 0.0017$], but no Phase*CP interaction [$\chi^2_{(1)} = 0.019$, $p = 0.89$]. Additionally, the overall mean FS was positive in SHAM mice but negative in GDX mice, confirmed by a main effect of GDX [$\chi^2_{(1)} = 5.088$, $p = 0.024$] (**Figure 2A**).

Females

Generalized estimating equation analysis of FS confirmed a significant main effect of CP [$\chi^2_{(1)} = 27.965$, $p < 0.000001$], but not of Phase [$\chi^2_{(1)} = 0.732$, $p = 0.39$] nor a Phase*CP interaction

[$\chi^2_{(1)} = 1.943$, $p = 0.16$], reflecting the fact that FS decreased from a positive value PRE to a negative value POST in both ACQ and REV (**Supplementary Figure 2B**). GEE also confirmed an X-dose*GDX*Phase interaction [$\chi^2_{(1)} = 4.821$, $p = 0.028$] (**Figure 2B**). However, the simple effects of X-dose (XX > XO in GDX in REV; $p = 0.046$), of GDX (GDX > SHAM in XO in ACQ; $p = 0.009$), and of Phase (ACQ > REV in GDX XO; $p = 0.020$) did not survive correction for multiple comparisons.

Maximum Number of Consecutive Incorrect Responses (MAXCI)

To quantify response perseveration, we calculated the MAXCI within each testing phase (ACQ scores served as a baseline from which the magnitude of the effect of reversal on perseveration was derived) (**Figures 3A,B**). This was an alternative way of quantifying perseveration as compared to methods based on the number of errors within sequential blocks of trials (e.g., Ragozzino et al., 2002). The MAXCI was chosen for two reasons: (1) Prior analysis of cumulative response records from this particular reversal task indicated that it is more sensitive to the magnitude of the initial bout of post-reversal perseveration and (2) it consistently shows no significant correlation with the complementary regressive error measure (**Figure 3D**).

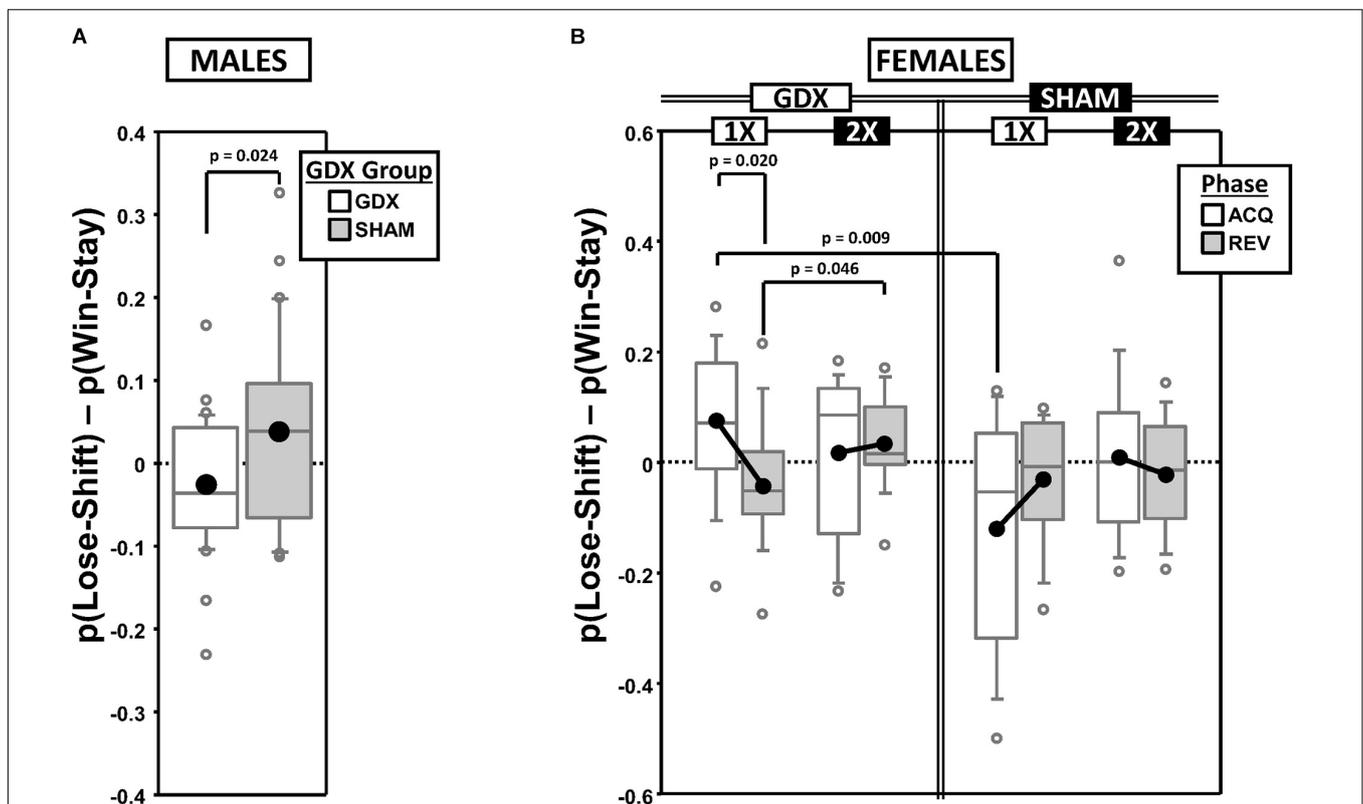
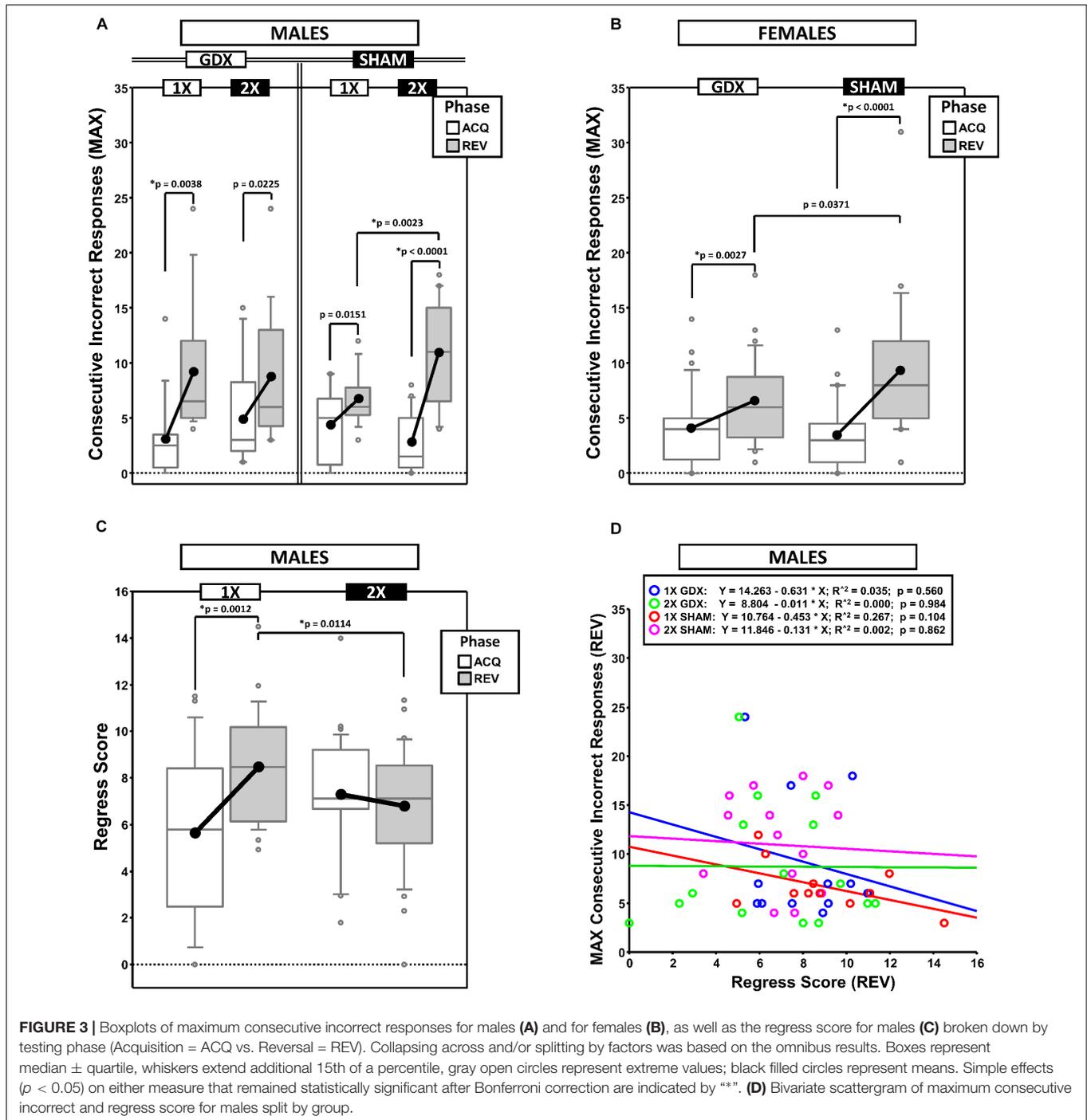


FIGURE 2 | Boxplots of the difference between the proportion of responses that were correct following either an incorrect response [$p(\text{Lose-Shift})$; index of flexibility] or a correct response [$p(\text{Win-Stay})$; index of stability] for males (**A**) broken down by gonadectomy group, or for females (**B**) broken down by X-dose (number of X-chromosomes; 1X or 2X), gonadectomy group and testing phase. Collapsing across and/or splitting by factors was based on the omnibus results. Boxes represent median \pm quartile, whiskers extend additional 15th of a percentile, gray open circles represent extreme values; black filled circles represent means. None of the simple effects ($p < 0.05$) indicated in (**B**) were significant after Bonferroni correction.



Males

Generalized estimating equation confirmed a main effect of Phase [$\chi^2_{(1)} = 45.719$, $p < 0.00001$] and a significant X-dose*GDX*Phase interaction [$\chi^2_{(1)} = 7.189$, $p = 0.0073$] (Figure 3A). *Post hoc* comparisons confirmed a simple effect of X-dose in the SHAM group in REV (XXY > XY; $p = 0.0023$; all other $p > 0.19$). By contrast, no simple effects of GDX were confirmed (all $p > 0.13$). Lastly, although MAXCI was greater in REV than ACQ for XY-GDX ($p = 0.0038$), XY-SHAM,

($p = 0.0151$), XXY-GDX ($p = 0.0225$), and XXY-SHAM ($p < 0.00001$), this difference was not significant for XY-SHAM and XXY-GDX after Bonferroni correction.

Females

Generalized estimating equation analysis confirmed a main effect of Phase [$\chi^2_{(1)} = 36.010$, $p < 0.00001$] and a significant GDX*Phase interaction [$\chi^2_{(1)} = 4.924$, $p = 0.027$]. Although the simple effect of Phase in GDX (REV > ACQ; $p = 0.0027$)

and SHAM (REV > ACQ; $p < 0.0001$) survived correction for multiple comparisons, the simple effect of GDX in REV did not (SHAM > GDX; $p = 0.0371$) (**Figure 3B**).

Regress Score (Regressive Errors on Path to Criterion)

To quantify the tendency for performance to initially improve and then regress away from criterion, we calculated the changes in accuracy across the phase (moving window: prior 5 trials – next 5 trials), summed the negative values and then normalized this sum to the total number of trials (**Figure 3C**).

Males

Generalized estimating equation analysis confirmed a main effect of Phase [$\chi^2_{(1)} = 4.792$, $p = 0.0286$], and an X-dose*Phase interaction [$\chi^2_{(1)} = 10.348$, $p = 0.0013$] (**Figure 3C**). *Post hoc* comparisons confirmed that Regress scores were higher in REV than ACQ in XY mice ($p = 0.0012$), but not XXY mice ($p = 0.35$). Additionally, the Regress score was higher in XY mice than XXY mice in REV ($p = 0.0114$), but not in ACQ ($p = 0.0517$).

Females

Generalized estimating equation analysis did not confirm any main effects or interactions (all $p > 0.14$).

Correlations: MAXCI, Regress Score, and Trials to Criterion

If MAXCI and the Regress Score quantify independent components of reversal learning (i.e., perseverative and regressive errors, respectively), then they should not be correlated. Correlational analyses split by sex, X-dose and GDX groups confirmed that these measures were not correlated (all $p > 0.1$, males illustrated in **Figure 3D**). However, both MAXCI and the Regress Score were positively correlated to trials to criterion (MAXCI, all $p < 0.008$; Regress Score, all $p < 0.007$). Thus, although errors of either type predicted trials to criterion, one error did not predict the other.

Premature Responding

Premature responses per trial to the target apertures (i.e., nose pokes into the unlit target apertures outside of the period of target presentation) on both the correct side in acquisition (C_{ACQ}) and the correct side in reversal (C_{REV}) were analyzed as a function of testing phase and change point (CP) (**Figures 4A–G**).

Males

Generalized estimating equation analysis of premature responding on the C_{ACQ} side confirmed a main effect of CP [$\chi^2_{(1)} = 7.905$, $p = 0.0049$] and a Phase*CP interaction [$\chi^2_{(1)} = 41.553$, $p < 0.00001$], but no main effect of Phase [$\chi^2_{(1)} = 0.008$, $p = 0.93$] (**Figure 4A**). However, neither X-dose, nor GDX, affected this pattern of results (all $p > 0.15$). The simple effects of Phase and Change point were all significant (all $p < 0.001$).

Generalized estimating equation analysis of premature responding on the C_{REV} side confirmed main effects of Phase [$\chi^2_{(1)} = 154.663$, $p < 0.00001$], CP [$\chi^2_{(1)} = 72.122$,

$p < 0.00001$], and a Phase*CP interaction [$\chi^2_{(1)} = 72.995$, $p < 0.00001$]. There was also an X-dose*GDX*Phase interaction [$\chi^2_{(1)} = 6.067$, $p = 0.0138$], X-dose*GDX*CP interaction [$\chi^2_{(1)} = 6.958$, $p = 0.0083$], X-dose*GDX*Phase*CP interaction [$\chi^2_{(1)} = 4.488$, $p = 0.034$], and a GDX*Phase*CP interaction [$\chi^2_{(1)} = 4.752$, $p = 0.029$] (**Figures 4C,D**). There were no significant simple effects of X-dose (all $p > 0.05$) and the simple effects of GDX (GDX > SHAM in XXY in both ACQ PRE and REV PRE; $p = 0.027$ and $p = 0.029$, respectively) did not survive correction for multiple comparisons (**Figure 4D**) (see **Figure 4C** for simple effects of Phase and CP).

Females

Generalized estimating equation analysis of premature responding on the C_{ACQ} side confirmed a main effect of CP [$\chi^2_{(1)} = 31.629$, $p < 0.00001$] and a Phase*CP interaction [$\chi^2_{(1)} = 77.525$, $p < 0.00001$], but no main effect of Phase [$\chi^2_{(1)} = 2.378$, $p = 0.12$] (**Figure 4B**; all simple effects of Phase and Change point significant; all $p < 0.0001$). GEE analysis also confirmed a main effect of X-dose [**Figure 4C**; $\chi^2_{(1)} = 4.506$, $p = 0.034$], with XX mice overall exhibiting more premature responses than XO subjects (**Figure 4F**). Finally, GEE confirmed a GDX*Phase*CP interaction [$\chi^2_{(1)} = 4.084$, $p = 0.043$], but none of the *post hoc* comparisons for the simple effect of GDX were significant after correction (all $p > 0.03$) (see **Figure 4E** for simple effects of Phase and CP).

Generalized estimating equation analysis of premature responding on the C_{REV} side confirmed main effects of Phase [$\chi^2_{(1)} = 160.374$, $p < 0.00001$] and CP [$\chi^2_{(1)} = 88.466$, $p < 0.00001$], as well as a Phase*CP interaction [$\chi^2_{(1)} = 86.810$, $p < 0.00001$; see **Figure 4B** for simple effects] and a GDX*Phase interaction [$\chi^2_{(1)} = 7.773$, $p = 0.0053$; **Figure 4G**]. Although *post hoc* comparisons confirmed the simple effect of phase within both GDX groups (both $p < 0.00001$), the simple effect of GDX in REV (GDX > SHAM, $p = 0.026$) did not survive Bonferroni correction.

Response Latencies

In addition to a decrease in response errors, successful learning of a spatial discrimination task leads to progressive decreases in response latencies. Thus, we calculated mean response latencies (i.e., time from target stimuli onset to first response to a target location) on both the correct side in acquisition (C_{ACQ}) and the correct side in reversal (C_{REV}) and analyzed these as a function of testing phase and CP (**Figures 5A,B** and **Supplementary Figures 3A,B**).

Males

Generalized estimating equation analysis of response latencies on the C_{ACQ} side confirmed main effects of Phase [$\chi^2_{(1)} = 16.226$, $p < 0.0001$] and CP [$\chi^2_{(1)} = 7.357$, $p = 0.007$], as well as a Phase*CP interaction [$\chi^2_{(1)} = 9.700$, $p = 0.002$]. However, on the C_{REV} side, analysis confirmed a main effect of Phase [$\chi^2_{(1)} = 18.521$, $p = 0.00002$], but no effect of CP [$\chi^2_{(1)} = 0.990$, $p = 0.3$] or Phase*CP interaction [$\chi^2_{(1)} = 1.465$, $p = 0.2$].

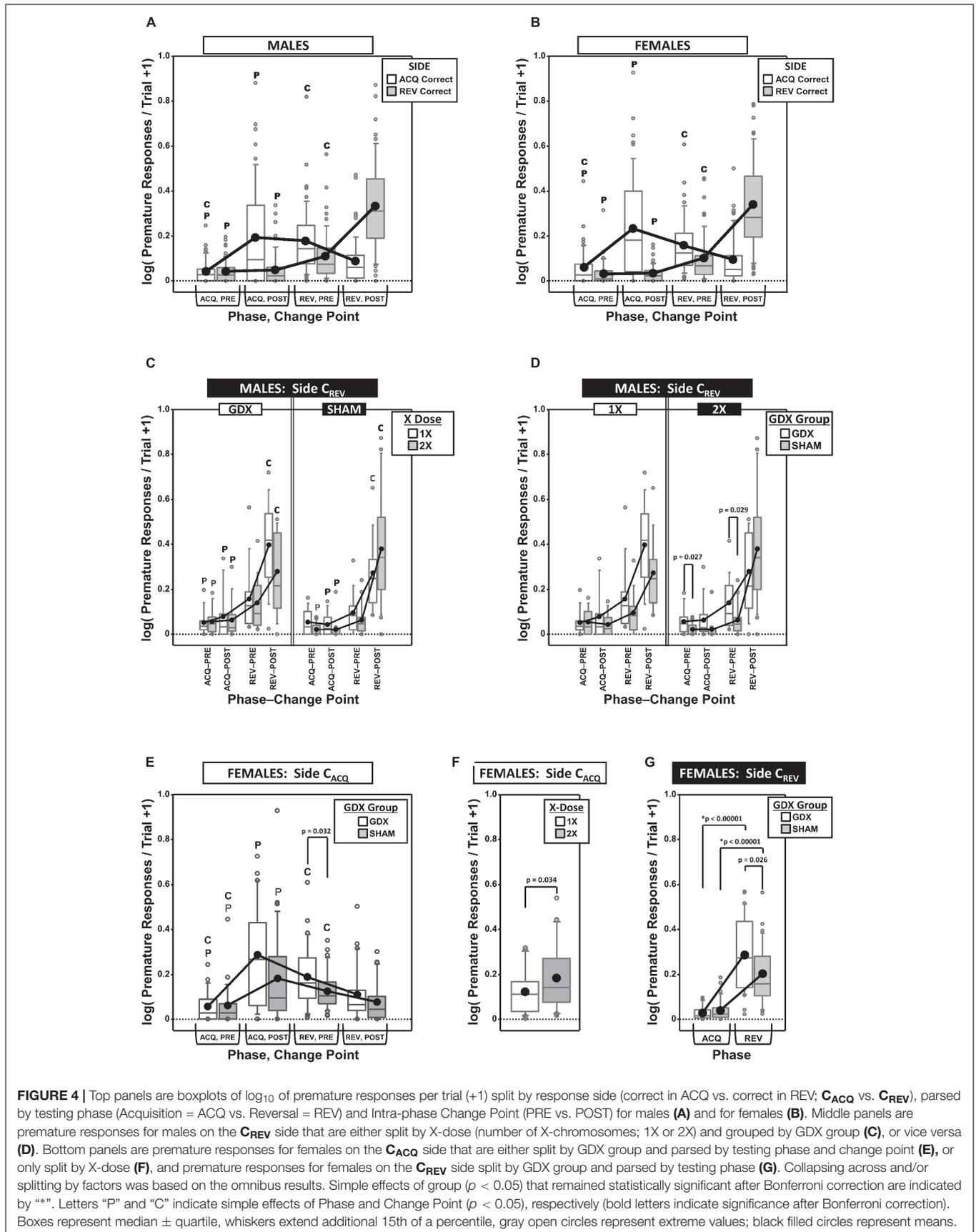


FIGURE 4 | Top panels are boxplots of log₁₀ of premature responses per trial (+1) split by response side (correct in ACQ vs. correct in REV; **C_{ACQ}** vs. **C_{REV}**), parsed by testing phase (Acquisition = ACQ vs. Reversal = REV) and Intra-phase Change Point (PRE vs. POST) for males (**A**) and for females (**B**). Middle panels are premature responses for males on the **C_{REV}** side that are either split by X-dose (number of X-chromosomes; 1X or 2X) and grouped by GDJ group (**C**), or vice versa (**D**). Bottom panels are premature responses for females on the **C_{ACQ}** side that are either split by GDJ group and parsed by testing phase and change point (**E**), or only split by X-dose (**F**), and premature responses for females on the **C_{REV}** side split by GDJ group and parsed by testing phase (**G**). Collapsing across and/or splitting by factors was based on the omnibus results. Simple effects of group ($p < 0.05$) that remained statistically significant after Bonferroni correction are indicated by “*”. Letters “P” and “C” indicate simple effects of Phase and Change Point ($p < 0.05$), respectively (bold letters indicate significance after Bonferroni correction). Boxes represent median \pm quartile, whiskers extend additional 15th of a percentile, gray open circles represent extreme values; black filled circles represent means.

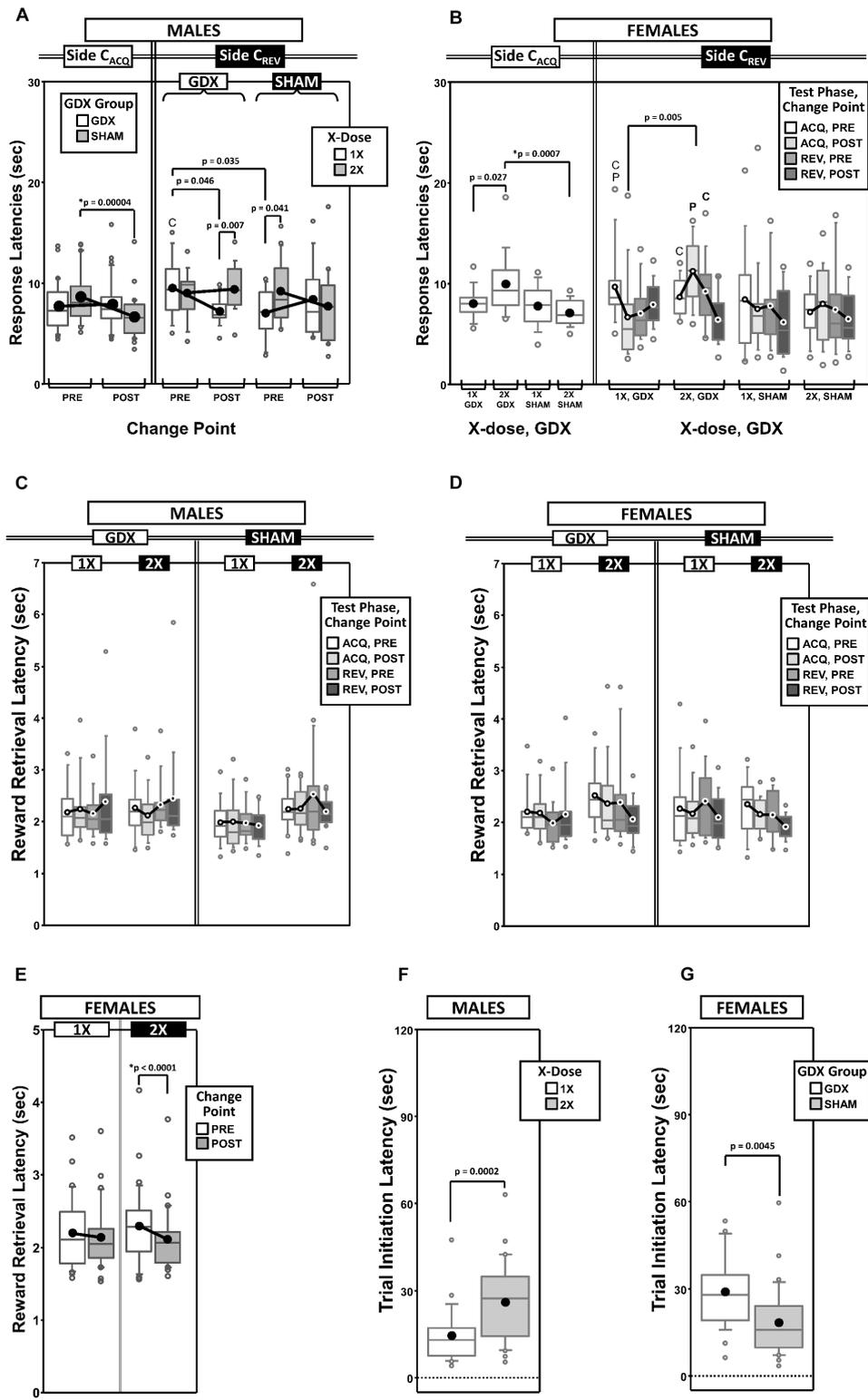


FIGURE 5 | Top panels are boxplots of response latencies to target apertures as a function of response side (correct in ACQ = C_{ACQ} vs. correct in REV = C_{REV}) for males (A) and for females (B), split by X-dose (number of X-chromosomes; 1X or 2X) or GDX group and parsed by testing phase (Acquisition = ACQ vs. Reversal = REV) and/or Intra-phase Change Point (PRE vs. POST). Middle panels are boxplots of reward-retrieval latencies for males (C) and for females (D) grouped by X-dose, GDX group, testing phase and CP. Bottom panels include box plots of reward retrieval latencies for females collapsed across testing phase and GDX (Continued)

FIGURE 5 | Continued

group (**E**) and of trial-initiation latencies for males split by X-dose (**F**) and for females split by GDX group (**G**). Collapsing across and/or splitting by factors was guided by the omnibus results. Simple effects of group ($p < 0.05$) on response latencies and trial initiation latencies that remained statistically significant after Bonferroni correction are indicated by “*”. Letters “P” and “C” indicate simple effects of Phase and Change Point ($p < 0.05$), respectively, on response latencies and trial initiation latencies (bold letters indicate significance after Bonferroni correction). On reward retrieval latencies in males, there was a significant GDX by Phase by Change point interaction ($p = 0.048$; no significant simple effects, all $p > 0.06$). On reward retrieval latencies in females, there were main effects of Phase ($p = 0.042$) and Change Point ($p = 0.011$), and an X-dose by Change Point interaction ($p = 0.033$; simple effect of Change Point in 2X, $p < 0.0001$; all other simple effects, $p > 0.12$). Boxes represent median \pm quartile, whiskers extend additional 15th of a percentile, gray open circles represent extreme values; filled circles represent means.

Response latencies to the C_{ACQ} side decreased during ACQ (POST CP $<$ PRE CP) but not REV (POST CP \approx PRE CP), while response latencies to the C_{REV} side decreased during REV (POST CP $<$ PRE CP) but not ACQ (POST CP \approx PRE CP) (**Supplementary Figure 3A**).

Analysis of the C_{ACQ} side also confirmed a significant GDX*CP interaction [$\chi^2_{(1)} = 11.732$, $p = 0.0006$]. *Post hoc* comparisons confirmed a simple effect of CP in SHAM mice ($p = 0.00004$) but not in GDX mice ($p = 0.6$); the simple effect of GDX was not significant either PRE ($p = 0.14$) or POST ($p = 0.07$) (**Figure 5A**).

Lastly, an X-dose*GDX*CP interaction was confirmed on the C_{REV} side [$\chi^2_{(1)} = 6.130$, $p = 0.013$], but *post hoc* comparisons did not identify any simple effects that survived correction for multiple comparisons (**Figure 5A**).

Females

Generalized estimating equation analysis of response latencies on the C_{ACQ} side confirmed main effects of Phase [$\chi^2_{(1)} = 22.846$, $p < 0.00001$] and CP [$\chi^2_{(1)} = 5.234$, $p = 0.02$], but no Phase*CP interaction [$\chi^2_{(1)} = 0.651$, $p = 0.4$] (**Supplementary Figure 3B**). A main effect of GDX [$\chi^2_{(1)} = 8.007$, $p = 0.005$], as well as an X-dose*GDX interaction [$\chi^2_{(1)} = 5.791$, $p = 0.016$] was identified (**Figure 5B**). *Post hoc* comparisons confirmed the simple effect of GDX in XX mice (SHAM $<$ GDX; $p = 0.0007$); however the simple effect of X-dose in GDX mice (1X $<$ 2X) was not significant after Bonferroni correction ($p = 0.027$) (other $p > 0.2$).

Generalized estimating equation analysis of response latencies on the C_{REV} side confirmed a main effect of Phase [$\chi^2_{(1)} = 5.996$, $p = 0.014$], but not of CP [$\chi^2_{(1)} = 3.019$, $p = 0.08$] nor a Phase*CP interaction [$\chi^2_{(1)} = 1.503$, $p = 0.2$] (**Supplementary Figure 3B**). GEE analysis also confirmed an X-dose*Phase*CP interaction [$\chi^2_{(1)} = 8.276$, $p = 0.004$] and an X-dose*GDX*Phase*CP interaction [$\chi^2_{(1)} = 4.361$, $p = 0.037$]. No simple effects of X-dose or GDX survived correction for multiple comparisons (**Figure 5B**).

Reward Retrieval Latencies

As an ancillary response latency measure, we calculated the average time to collect the reward after committing a correct response as a function of both testing phase and CP (**Figure 5C** thru **Figure 5D**).

Males

Generalized estimating equation analysis of reward retrieval latencies confirmed no main effects of Phase [$\chi^2_{(1)} = 0.468$, $p = 0.5$] or CP [$\chi^2_{(1)} = 0.130$, $p = 0.7$], nor a Phase*CP interaction [$\chi^2_{(1)} = 0.059$, $p = 0.8$]. GEE analysis did confirm

a GDX*Phase*CP interaction [$\chi^2_{(1)} = 3.914$, $p = 0.048$], however none of the *post hoc* comparisons were significant (all $p > 0.06$) (**Figure 5C**).

Females

Generalized estimating equation analysis of reward retrieval latencies did confirm main effects of Phase [$\chi^2_{(1)} = 4.151$, $p = 0.04$; ACQ $>$ REV] and CP [$\chi^2_{(1)} = 6.495$, $p = 0.01$; PRE $>$ POST], but no Phase*CP interaction [$\chi^2_{(1)} = 0.257$, $p = 0.6$] (**Figure 5D**). GEE analysis also confirmed an X-dose*CP interaction [$\chi^2_{(1)} = 4.539$, $p = 0.03$]. *Post hoc* comparisons confirmed a simple effect of CP in XX mice ($p < 0.0001$) (all other $p > 0.1$) (**Figure 5E**).

Trial Initiation Latencies

As an ancillary response latency measure that is affected by motivation to engage in the task, we calculated the mean time to initiate a trial (i.e., time from the end of inter-trial interval to the 1st observing response) as a function of both testing phase and the outcome of the prior trial (rewarded or unrewarded) (**Figures 5F,G** and **Supplementary Figures 3C,D**).

Males

As expected, trial initiation latencies were shorter after unrewarded trials ($M = 12.0$ s, $SD = 8.3$ s) than after rewarded trials ($M = 29.5$ s, $SD = 20.1$ s), regardless of phase (**Supplementary Figure 3C**). GEE analysis on latencies confirmed main effects of Phase [$\chi^2_{(1)} = 11.387$, $p = 0.0007$] and the outcome of the prior trial [$\chi^2_{(1)} = 134.876$, $p < 0.00001$], but no Phase*Prior-Trial interaction [$\chi^2_{(1)} = 1.473$, $p = 0.2$]. GEE also confirmed a main effect of X-dose [$\chi^2_{(1)} = 14.300$, $p = 0.00016$]. XXY mice had overall longer trial initiation latencies than XY mice (**Figure 5F**).

Females

As expected, trial initiation latencies were shorter after unrewarded trials ($M = 14.4$ s, $SD = 9.8$ s) than rewarded trials ($M = 33.3$ s, $SD = 18.2$ s) (**Supplementary Figure 3D**). GEE analysis on latencies confirmed main effects of Phase [$\chi^2_{(1)} = 4.379$, $p = 0.036$] and the outcome of the prior trial [$\chi^2_{(1)} = 151.134$, $p < 0.00001$], but no Phase*Prior-Trial interaction [$\chi^2_{(1)} = 1.086$, $p = 0.3$]. GEE also confirmed a main effect of GDX [$\chi^2_{(1)} = 8.088$, $p = 0.0045$]. GDX mice had overall longer latencies than SHAM mice (**Figure 5G**).

Timeouts From Observing-Response Failures (ORTO)

Failure to complete the holding requirement of the observing response (~0.0–0.2 s) causes a short 2-s timeout after which the subject must make another attempt; thus, a subject can incur multiple observing-response timeouts each trial. These errors may reflect a facet of response inhibition or impulsivity (Linden et al., 2018); thus, we calculated the mean number of timeouts per trial as a function of both testing phase and CP (Figures 6A,B).

Males

Generalized estimating equation analysis of ORTO confirmed a main effect of Phase [REV > ACQ; $\chi^2_{(1)} = 18.311, p = 0.0001$], but no main effect of CP [$\chi^2_{(1)} = 2.089, p = 0.15$], nor a Phase*CP interaction [$\chi^2_{(1)} = 0.611, p = 0.4$]. GEE analysis also confirmed a GDX*Phase*CP interaction [$\chi^2_{(1)} = 4.263, p = 0.039$] and an X-dose*GDX*Phase*CP interaction [$\chi^2_{(1)} = 4.780, p = 0.029$]. However, none of the *post hoc* comparisons were significant after Bonferroni correction (Figure 6A).

Females

Generalized estimating equation analysis of ORTO confirmed a main effect of CP [$\chi^2_{(1)} = 10.091, p = 0.0015$] and a Phase*CP interaction [$\chi^2_{(1)} = 4.075, p = 0.044$], but no main effect of Phase [$\chi^2_{(1)} = 3.103, p = 0.078$] (Figure 6B). *Post hoc* comparisons confirmed a simple effect of Phase PRE (REV > ACQ; $p = 0.01$) and a simple effect of CP in ACQ (POST > PRE; $p = 0.00005$) (other $p > 0.3$). There were no effects of, or interactions with, either X-dose or GDX (all $p \geq 0.1$).

Extraneous Observing Responses (EOR)

During the period of target aperture illumination, observing responses were without consequence. However, as these EOR may reflect a lack of attention and/or the degree to which the subject has learned the task rules (EOR are expected to decrease and the animal gains proficiency), we calculated the mean EOR per trial parsed by Phase and CP (Figures 6C,D).

Males

GEE analysis of EOR confirmed a main effect of Phase ($\chi^2_{(1)} = 14.598, p = 0.0001$) and GDX ($\chi^2_{(1)} = 6.018, p = 0.014$), but no main effect of CP ($\chi^2_{(1)} = 0.023, p = 0.9$) nor Phase*CP interaction ($\chi^2_{(1)} = 0.076, p = 0.8$). EOR decreased from ACQ to REV and were higher in GDX mice than SHAM mice (Figure 6C).

Females

Generalized estimating equation analysis of EOR confirmed a main effect of Phase [ACQ > REV; $\chi^2_{(1)} = 17.410, p = 0.00003$] and GDX [$\chi^2_{(1)} = 4.647, p = 0.03$], but no main effect of CP [$\chi^2_{(1)} = 0.724, p = 0.4$] nor Phase*CP interaction [$\chi^2_{(1)} = 0.971, p = 0.3$]. GEE analysis also confirmed an X-dose*Phase interaction [$\chi^2_{(1)} = 5.013, p = 0.025$] and a GDX*Phase interaction [$\chi^2_{(1)} = 6.296, p = 0.012$]. *Post hoc* comparisons confirmed simple effects of Phase in XX mice ($p < 0.00001$) and GDX mice ($p < 0.00001$), as well as a simple

effect of GDX in ACQ (GDX > SHAM; $p = 0.005$) (all other $p > 0.08$) (Figure 6D).

Omissions

Failure to make a response into a target aperture during the 30-s target-stimuli presentation interval resulted in a 5-s timeout and the recording of an omission error. We calculated the proportion of trials that were omission errors parsed by testing phase and CP (Figures 7A,B and Supplementary Figures 4A,B).

Males

Generalized estimating equation analysis of omission errors confirmed a main effect of X-dose [$\chi^2_{(1)} = 4.452, p = 0.035$], Phase [$\chi^2_{(1)} = 16.845, p = 0.00004$], and CP [$\chi^2_{(1)} = 67.598, p < 0.00001$], as well as a Phase*CP interaction [$\chi^2_{(1)} = 12.507, p = 0.0004$]. As expected, omission errors decreased from ACQ-PRE to REV-POST (Supplementary Figure 4A). Omission errors were overall more frequent in XXY than XY mice (Figure 7A).

Females

Generalized estimating equation analysis of omission errors confirmed the main effects of Phase [$\chi^2_{(1)} = 10.543, p = 0.001$] and CP [$\chi^2_{(1)} = 68.951, p < 0.00001$], but no Phase*CP interaction [$\chi^2_{(1)} = 2.164, p = 0.14$]. As expected, omission errors decreased from ACQ-PRE to REV-POST (Supplementary Figure 4B). GEE analysis also confirmed an X-dose*GDX interaction [$\chi^2_{(1)} = 6.671, p = 0.0098$]. *Post hoc* comparisons confirmed simple effects of X-dose in GDX mice (XX > XO; $p = 0.0027$) and of GDX in XX mice (GDX > SHAM; $p = 0.0009$) (other $p > 0.4$). Omissions were overall more frequent in XX-GDX mice than in the other groups (Figure 7B).

DISCUSSION

Overview

In gonadally intact males, XXY mice exhibited a greater degree of perseverative error in the reversal phase than XY mice, while in gonadectomized males, XXY mice required more trials to reach criterion in the acquisition phase than XY mice. These results indicate that gonadal function moderates the impact of X-dose on impulsivity and learning, respectively, in the XY* model of KS. Nevertheless, in KS men, androgen therapy in adulthood does not appear to affect these aspects of learning and cognition (Fales et al., 2003; Kompus et al., 2011; Liberato et al., 2017; Skakkebaek et al., 2017).

Additionally, a number of gonadectomy-independent effects were observed (overall, longer trial initiation latencies and more omissions by XXY compared to XY; greater regressive error in reversal in XY than XXY) that indicate direct effects of the sex-chromosome complement on behavior, even if the relevance of these effects to KS is currently either unclear or unsubstantiated.

Consistent with past reports (Davies et al., 2005, 2007), XO vs. XX comparisons of RLT performance do not model the cognitive deficits in TS.

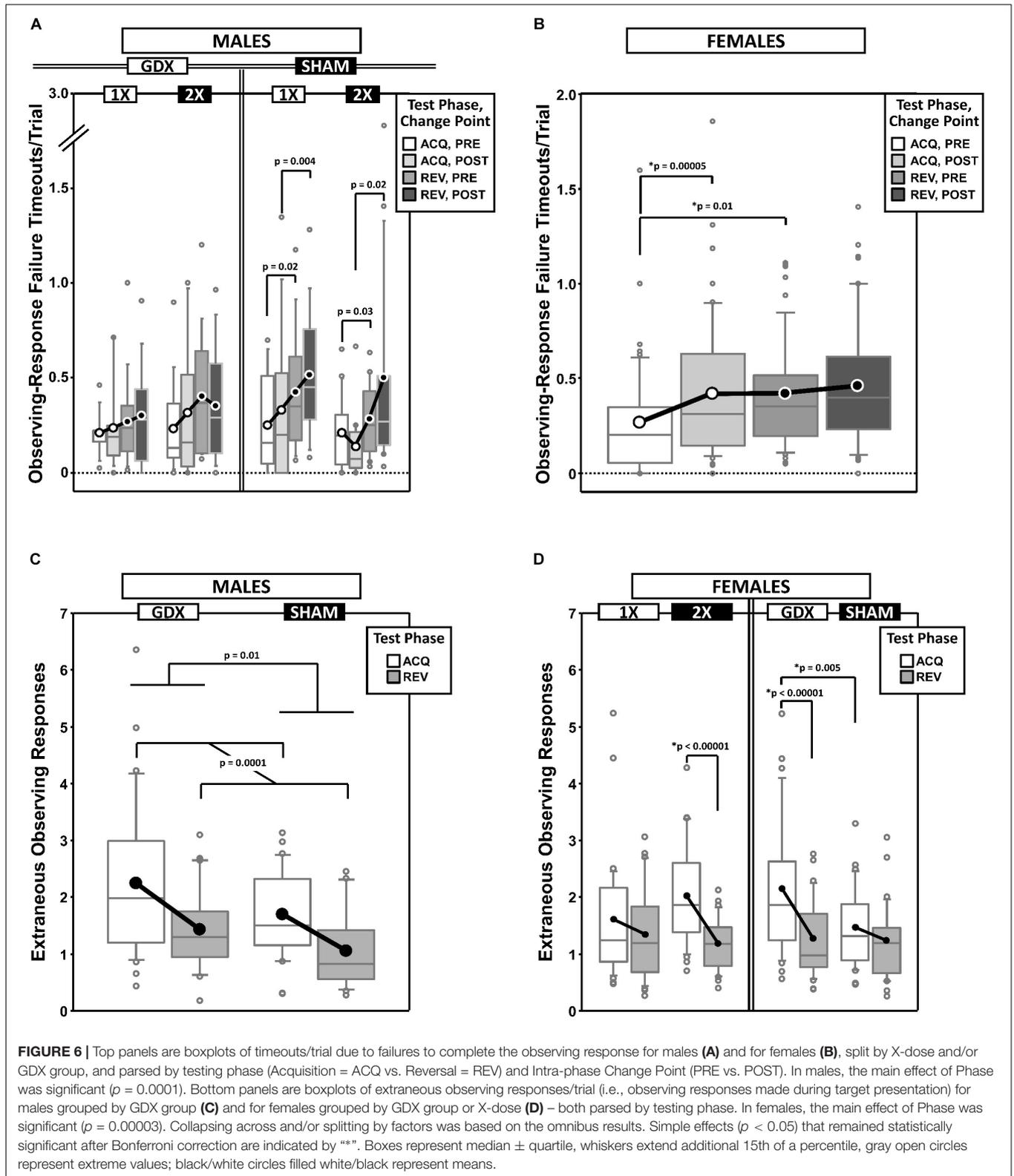
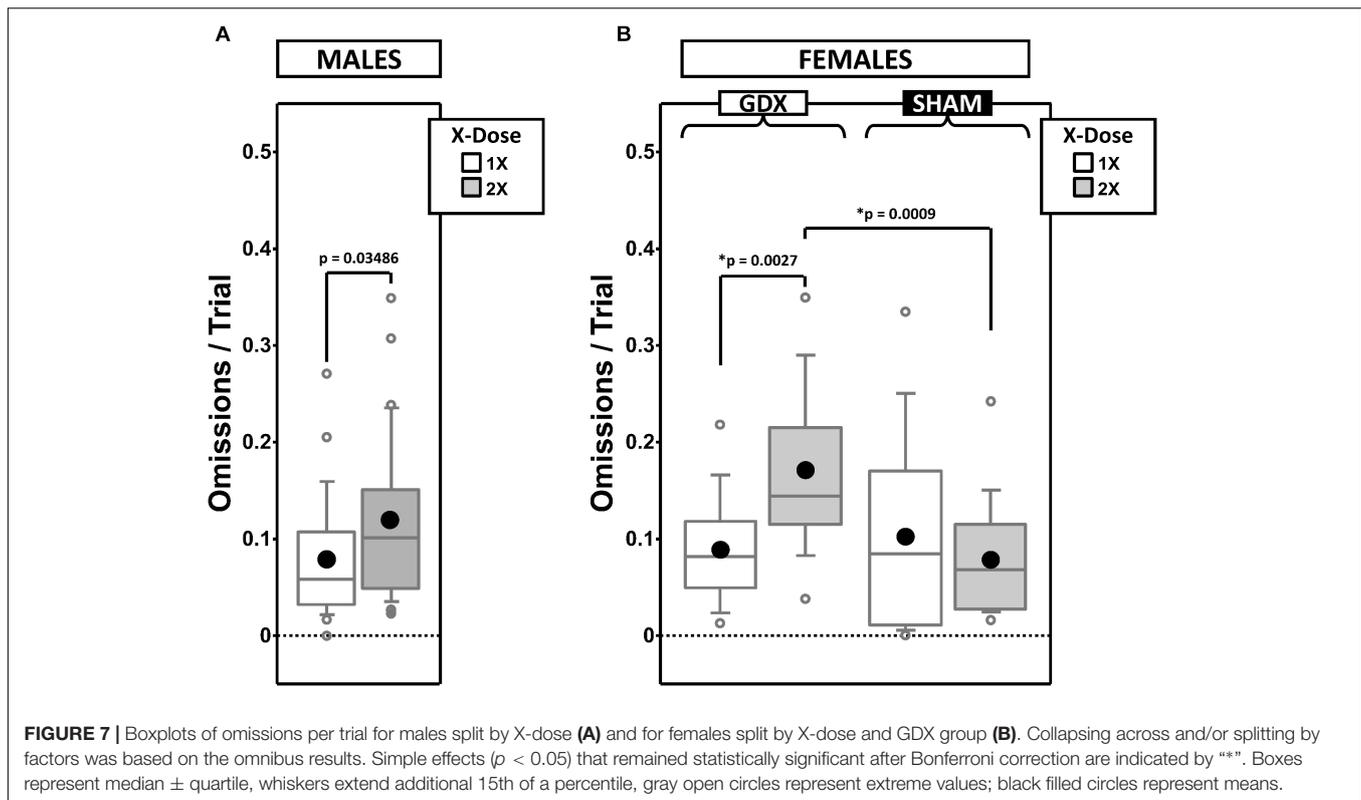


FIGURE 6 | Top panels are boxplots of timeouts/trial due to failures to complete the observing response for males (A) and for females (B), split by X-dose and/or GDX group, and parsed by testing phase (Acquisition = ACQ vs. Reversal = REV) and Intra-phase Change Point (PRE vs. POST). In males, the main effect of Phase was significant ($p = 0.0001$). Bottom panels are boxplots of extraneous observing responses/trial (i.e., observing responses made during target presentation) for males grouped by GDX group (C) and for females grouped by GDX group or X-dose (D) – both parsed by testing phase. In females, the main effect of Phase was significant ($p = 0.00003$). Collapsing across and/or splitting by factors was based on the omnibus results. Simple effects ($p < 0.05$) that remained statistically significant after Bonferroni correction are indicated by “*”. Boxes represent median \pm quartile, whiskers extend additional 15th of a percentile, gray open circles represent extreme values; black/white circles filled white/black represent means.

XXY vs. XY Mice as a Model of KS

Two principal measures discriminated XXY from XY mice – trials to criterion and perseverative responding in reversal – and

both of these interacted with gonadal status. The group difference in perseverative responding is noteworthy for three reasons: (1) the greater perseverative responding in XXY than XY mice was



observed in intact mice only, (2) this effect is conceptually similar to the relatively higher number of perseverative errors made by people with KS on the Wisconsin Card Sorting Test (van Rijn et al., 2012a; Skakkebaek et al., 2013), and (3) both XXY mice (Lue et al., 2005; Lewejohann et al., 2009) and men with KS are hypogonadal, so this result may be due to a group difference in levels of gonadal hormones.

Additionally, these differences appear to be specific increases in perseverative responding, as overall error rates (errors/trial) did not differ between genotypes. Moreover, although XY mice showed a greater across-phase increase in, and higher reversal-phase values for, *regressive* errors than XXY mice, *regressive* errors did not correlate with perseverative errors. Additionally, van Rijn et al. (2012a) demonstrated that perseveration errors in KS men were not affected by androgen supplementation. Thus, resolving the significance for model validity of this genotype-by-gonadectomy interaction on perseverative errors will require further study.

Similarly, the implications of the gonadectomy-dependent learning deficit in XXY mice – but, not in XY mice – remains unclear. In van Rijn et al. (2012a), perseveration errors in KS men were statistically independent of overall WCST performance and intellectual function (full-scale IQ). However, in the RLT, perseverative errors were positively correlated with trials to criterion. Moreover, the pattern of group differences in trials to criterion did not mirror those for perseverative errors. As these observations place further limitations on model validity, it appears that alternative behavioral assays of perseverative responding and learning rate, as well as additional

controls for hormone levels/exposure, would be required to fully determine model validity.

XO vs. XX Mice as a Model of TS

Consistent with past reports comparing 40,XO mice to XX mice on an outbred MF1 strain (Davies et al., 2007), when comparing 40,XO mice to XX mice on an inbred C57BL/6J background in the RLT, 40,XO do not show cognitive deficits. Indeed, XX mice made more premature responses than XO mice on the initially reinforced side – a result opposite of that predicted for a model of TS – and no genotypic effect was observed on the side reinforced in the reversal phase. Additionally, there were no significant effects of genotype on either perseverative responding or response-holding failures (impulsivity measure), and the genotype effects on response latencies did not survive statistical correction. Lastly, although the across-phase decrease in EORs (attention measure) was statistically significant in XX but not XO mice, within-phase genotypic differences were not.

Thus, although the RLT was sufficiently sensitive to detect a number of gonadectomy effects on these measures (larger reversal cost on perseverative errors in SHAM than GDX, more premature responding in GDX than SHAM, longer response latencies in XX-GDX than XX-SHAM, and more acquisition phase EORs in GDX than SHAM), differences caused by genotype were either below the limit of detection or opposite of that predicted for a model of TS. However, these results do bolster the claim that the small Y^{*X} chromosome in the 40,XO mice can largely rescue the cognitive deficits that may have resulted from the complete loss of a second X chromosome. That suggestion

could increase the attractiveness of the steroid sulfatase gene (*STS*) as a target for TS-associated cognitive deficits (Davies et al., 2005, 2007, 2009).

A difference in copy number of *Sts* does not explain the differences between XY and XXY mice reported here as both have only one copy of *Sts* (both Y* and X^{Y*} lack *Sts*) (Burgoyne and Arnold, 2016). Nevertheless, dysregulation of *Sts* in 40, XXY mice may still occur (e.g., due to skewed X-inactivation – X vs. X^{Y*} – and/or escape from inactivation). Interest in *STS* as a contributor to phenotypic features of KS remains high because of previous studies in mice (Davies et al., 2009, 2014), and because of the association of *STS* with cognitive function and deficits in humans (Kent et al., 2008; Stergiakouli et al., 2011; Chatterjee et al., 2016; Cavenagh et al., 2019).

CONCLUSION

Greater perseveration in XXY than XY intact male mice in the RLT is a facet of inflexible responding with face validity for the greater perseverative responding in KS men than controls on the WCST. Importantly, the effects of sex hormones appeared to play a moderating role, as this genotypic difference was not observed in GDX mice. These results may have implications for the use of androgens to remedy the cognitive deficits observed in KS men. Although androgen therapy does not appear to improve cognition in either KS adults (Skakkebaek et al., 2017) or children (Ross et al., 2017), little is known about the consequence of either an increased severity of hypogonadism and/or the lack of standard androgen replacement therapy on cognitive function. Our results indicate that in XXY mice, even small changes in circulating androgens can significantly change cognitive processes and learning.

REFERENCES

- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., and Clubley, E. (2001). The autism-spectrum quotient (AQ): evidence from asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *J. Autism Dev. Disord.* 31, 5–17.
- Belling, K., Russo, F., Jensen, A. B., Dalgaard, M. D., Westergaard, D., Rajpert-De Meyts, E., et al. (2017). Klinefelter syndrome comorbidities linked to increased X chromosome gene dosage and altered protein interactome activity. *Hum. Mol. Genet.* 26, 1219–1229. doi: 10.1093/hmg/ddx014
- Burgoyne, P. S., and Arnold, A. P. (2016). A primer on the use of mouse models for identifying direct sex chromosome effects that cause sex differences in non-gonadal tissues. *Biol. Sex Differ.* 7:68. doi: 10.1186/s13293-016-0115-5
- Burgoyne, P. S., Mahadevaiah, S. K., Perry, J., Palmer, S. J., and Ashworth, A. (1998). The Y* rearrangement in mice: new insights into a perplexing PAR. *Cytogenet. Cell Genet.* 80, 37–40. doi: 10.1159/000014954
- Cavenagh, A., Chatterjee, S., and Davies, W. (2019). Behavioural and psychiatric phenotypes in female carriers of genetic mutations associated with X-linked ichthyosis. *PLoS One* 14:e0212330. doi: 10.1371/journal.pone.0212330
- Cederlöf, M., Ohlsson Gotby, A., Larsson, H., Serlachius, E., Boman, M., Långström, N., et al. (2014). Klinefelter syndrome and risk of psychosis, autism and ADHD. *J. Psychiatr. Res.* 48, 128–130. doi: 10.1016/j.jpsychires.2013.10.001
- Chatterjee, S., Humby, T., and Davies, W. (2016). Behavioural and psychiatric phenotypes in men and boys with X-linked ichthyosis: evidence from a worldwide online survey. *PLoS One* 11:e0164417. doi: 10.1371/journal.pone.0164417

DATA AVAILABILITY

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

Animal Subjects: The animal study was reviewed and approved by the UCLA Chancellor's Animal Research Committee.

AUTHOR CONTRIBUTIONS

SA, AA, and JJ designed the study and revised the manuscript. SA and HH conducted the study. SA collected and analyzed the data, and wrote the first draft of the manuscript. AA and JJ wrote the sections of the manuscript. SA, AA, and JJ revised the manuscript.

FUNDING

Funding for this study was provided by the NIH grant HD076125.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnbeh.2019.00201/full#supplementary-material>

- Chen, X., McClusky, R., Itoh, Y., Reue, K., and Arnold, A. P. (2013a). X and Y chromosome complement influence adiposity and metabolism in mice. *Endocrinology* 154, 1092–1104. doi: 10.1210/en.2012-98
- Chen, X., Williams-Burris, S. M., McClusky, R., Ngun, T. C., Ghahramani, N., Barseghyan, H., et al. (2013b). The sex chromosome trisomy mouse model of XXY and XYY: metabolism and motor performance. *Biol. Sex Differ.* 4:15. doi: 10.1186/2042-6410-4-15
- Cox, K. H., Bonthuis, P. J., and Rissman, E. F. (2014). Mouse model systems to study sex chromosome genes and behavior: relevance to humans. *Front. Neuroendocrinol.* 35:405–419. doi: 10.1016/j.yfrne.2013.12.004
- Cox, K. H., Quinnes, K. M., Eschendoeder, A., Didrick, P. M., Eugster, E. A., and Rissman, E. F. (2015). Number of X-chromosome genes influences social behavior and vasopressin gene expression in mice. *Psychoneuroendocrinology* 51, 271–281. doi: 10.1016/j.psyneuen.2014.10.010
- Davies, W., Humby, T., Isles, A. R., Burgoyne, P. S., and Wilkinson, L. S. (2007). X-monosomy effects on visuospatial attention in mice: a candidate gene and implications for Turner syndrome and attention deficit hyperactivity disorder. *Biol. Psychiatry* 61, 1351–1360. doi: 10.1016/j.biopsych.2006.08.011
- Davies, W., Humby, T., Kong, W., Otter, T., Burgoyne, P. S., and Wilkinson, L. S. (2009). Converging pharmacological and genetic evidence indicates a role for steroid sulfatase in attention. *Biol. Psychiatry* 66, 360–367. doi: 10.1016/j.biopsych.2009.01.001
- Davies, W., Humby, T., Trent, S., Eddy, J. B., Ojarikre, O. A., and Wilkinson, L. S. (2014). Genetic and pharmacological modulation of the steroid sulfatase axis improves response control; comparison with drugs used in ADHD. *Neuropsychopharmacology* 39, 2622–2632. doi: 10.1038/npp.2014.115

- Davies, W., Isles, A., Smith, R., Karunadasa, D., Burrmann, D., Humby, T., et al. (2005). Xlr3b is a new imprinted candidate for X-linked parent-of-origin effects on cognitive function in mice. *Nat. Genet.* 37, 625–629. doi: 10.1038/ng1577
- Deloof, J., Van den Berghe, H., Swillen, A., Kleczkowska, A., and Fryns, J. P. (1993). Turner syndrome patients as adults: a study of their cognitive profile, psychosocial functioning and psychopathological findings. *Genet. Couns.* 4, 169–179.
- Eicher, E. M., Hale, D. W., Hunt, P. A., Lee, B. K., Tucker, P. K., King, T. R., et al. (1991). The mouse Y* chromosome involves a complex rearrangement, including interstitial positioning of the pseudoautosomal region. *Cytogenet. Cell Genet.* 57, 221–230. doi: 10.1159/000133152
- Fales, C. L., Knowlton, B. J., Holyoak, K. J., Geschwind, D. H., Swerdloff, R. S., and Gonzalo, I. G. (2003). Working memory and relational reasoning in Klinefelter syndrome. *J. Int. Neuropsychol. Soc.* 9, 839–846. doi: 10.1017/S1355617703960036
- Gallistel, C. R., Fairhurst, S., and Balsam, P. (2004). The learning curve: implications of a quantitative analysis. *Proc. Natl. Acad. Sci. U.S.A.* 101, 13124–13131. doi: 10.1073/pnas.0404965101
- Garron, D. C. (1977). Intelligence among persons with Turner's syndrome. *Behav. Genet.* 7, 105–127. doi: 10.1007/BF01066000
- Gravholt, C. H., Chang, S., Wallentin, M., Fedder, J., Moore, P., and Skakkebaek, A. (2018). Klinefelter syndrome: integrating genetics. *Neuropsychol. Endocrinol. Endocr. Rev.* 39, 389–423. doi: 10.1210/er.2017-212
- Green, T., Sagar, M., Ishak, A., Hong, D. S., and Reiss, A. L. (2018). X-Chromosome effects on attention networks: insights from imaging resting-state networks in turner syndrome. *Cereb. Cortex N. Y. N* 1991, 3176–3183. doi: 10.1093/cercor/bhx188
- Hong, D. S., and Reiss, A. L. (2014). Cognitive and neurological aspects of sex chromosome aneuploidies. *Lancet Neurol.* 13, 306–318. doi: 10.1016/S1474-4422(13)70302-8
- Isles, A. R., Davies, W., Burrmann, D., Burgoyne, P. S., and Wilkinson, L. S. (2004). Effects on fear reactivity in XO mice are due to haploinsufficiency of a non-PAR X gene: implications for emotional function in Turner's syndrome. *Hum. Mol. Genet.* 13, 1849–1855. doi: 10.1093/hmg/ddh203
- Kent, L., Emerton, J., Bhadravathi, V., Weisblatt, E., Pasco, G., Willatt, L. R., et al. (2008). X-linked ichthyosis (steroid sulfatase deficiency) is associated with increased risk of attention deficit hyperactivity disorder, autism and social communication deficits. *J. Med. Genet.* 45, 519–524. doi: 10.1136/jmg.2008.057729
- Klanker, M., Sandberg, T., Joosten, R., Willuhn, I., Feenstra, M., and Denys, D. (2015). Phasic dopamine release induced by positive feedback predicts individual differences in reversal learning. *Neurobiol. Learn. Mem.* 125, 135–145. doi: 10.1016/j.nlm.2015.08.011
- Klein, K. O., Rosenfield, R. L., Santen, R. J., Gawlik, A. M., Backeljauw, P. F., Gravholt, C. H., et al. (2018). Estrogen replacement in turner syndrome: literature review and practical considerations. *J. Clin. Endocrinol. Metab.* 103, 1790–1803. doi: 10.1210/je.2017-02183
- Kompus, K., Westerhausen, R., Nilsson, L.-G., Hugdahl, K., Jongstra, S., Berglund, A., et al. (2011). Deficits in inhibitory executive functions in klinefelter (47, XXY) syndrome. *Psychiatry Res.* 189, 135–140. doi: 10.1016/j.psychres.2011.02.028
- Laughlin, R. E., Grant, T. L., Williams, R. W., and Jentsch, J. D. (2011). Genetic dissection of behavioral flexibility: reversal learning in mice. *Biol. Psychiatry* 69, 1109–1116. doi: 10.1016/j.biopsych.2011.01.014
- Lepage, J.-F., Hong, D. S., Hallmayer, J., and Reiss, A. L. (2012). Genomic imprinting effects on cognitive and social abilities in prepubertal girls with turner syndrome. *J. Clin. Endocrinol. Metab.* 97, E460–E464. doi: 10.1210/jc.2011-916
- Lewejohann, L., Damm, O. S., Luetjens, C. M., Hämäläinen, T., Simoni, M., Nieschlag, E., et al. (2009). Impaired recognition memory in male mice with a supernumerary X chromosome. *Physiol. Behav.* 96, 23–29. doi: 10.1016/j.physbeh.2008.08.007
- Liberato, D., Granato, S., Grimaldi, D., Rossi, F. M., Tahani, N., Gianfrilli, D., et al. (2017). Fluid intelligence, traits of personality and personality disorders in a cohort of adult KS patients with the classic 47, XXY karyotype. *J. Endocrinol. Invest.* 40, 1191–1199. doi: 10.1007/s40618-017-0674-2
- Linden, J., James, A. S., McDaniel, C., and Jentsch, J. D. (2018). Dopamine D2 receptors in dopaminergic neurons modulate performance in a reversal learning task in mice. *eNeuro* 5:ENEURO.229-17.2018. doi: 10.1523/ENEURO.0229-17.2018
- Loesch, D. Z., Bui, Q. M., Kelso, W., Huggins, R. M., Slater, H., Warne, G., et al. (2005). Effect of Turner's syndrome and X-linked imprinting on cognitive status: analysis based on pedigree data. *Brain Dev.* 27, 494–503. doi: 10.1016/j.braindev.2004.12.009
- Lopes, A. M., Burgoyne, P. S., Ojarikre, A., Bauer, J., Sargent, C. A., Amorim, A., et al. (2010). Transcriptional changes in response to X chromosome dosage in the mouse: implications for X inactivation and the molecular basis of Turner Syndrome. *BMC Genomics* 11:82. doi: 10.1186/1471-2164-11-82
- Lue, Y., Jentsch, J. D., Wang, C., Rao, P. N., Hikim, A. P. S., Salameh, W., et al. (2005). XXY mice exhibit gonadal and behavioral phenotypes similar to Klinefelter syndrome. *Endocrinology* 146, 4148–4154. doi: 10.1210/en.2005-278
- Lue, Y., Wang, C., Liu, P. Y., Erkilli, K., and Swerdloff, R. S. (2010). Insights into the pathogenesis of XXY phenotype from comparison of the clinical syndrome with an experimental XXY mouse model. *Pediatr. Endocrinol. Rev.* 8(Suppl. 1), 140–144.
- Mauger, C., Lancelot, C., Roy, A., Coutant, R., Cantisano, N., and Le Gall, D. (2018). Executive functions in children and adolescents with turner syndrome: a systematic review and meta-analysis. *Neuropsychol. Rev.* 28, 188–215. doi: 10.1007/s11065-018-9372-x
- Mazzocco, M. M. M. (2006). The cognitive phenotype of Turner syndrome: specific learning disabilities. *Int. Congr. Ser.* 1298, 83–92. doi: 10.1016/j.ics.2006.06.016
- McCauley, E., Kay, T., Ito, J., and Treder, R. (1987). The turner syndrome: cognitive deficits, affective discrimination, and behavior problems. *Child Dev.* 58, 464–473. doi: 10.2307/1130523
- Papachristos, E. B., and Gallistel, C. R. (2006). Autoshapec head poking in the mouse: a quantitative analysis of the learning curve. *J. Exp. Anal. Behav.* 85, 293–308. doi: 10.1901/jeab.2006.71-5
- Ragozzino, M. E., Jih, J., and Tzavos, A. (2002). Involvement of the dorsomedial striatum in behavioral flexibility: role of muscarinic cholinergic receptors. *Brain Res.* 953, 205–214. doi: 10.1016/S0006-8993(02)03287-0
- Raznahan, A., Probst, F., Palmert, M. R., Giedd, J. N., and Lerch, J. P. (2013). High resolution whole brain imaging of anatomical variation in XO, XX, and XY mice. *NeuroImage* 83, 962–968. doi: 10.1016/j.neuroimage.2013.07.052
- Romans, S. M., Roeltgen, D. P., Kushner, H., and Ross, J. L. (1997). Executive function in girls with turner's syndrome. *Dev. Neuropsychol.* 13, 23–40. doi: 10.1080/87565649709540666
- Romans, S. M., Stefanatos, G., Roeltgen, D. P., Kushner, H., and Ross, J. L. (1998). Transition to young adulthood in Ullrich-Turner syndrome: neurodevelopmental changes. *Am. J. Med. Genet.* 79, 140–147.
- Ross, J. L., Kushner, H., Kowal, K., Bardsley, M., Davis, S., Reiss, A. L., et al. (2017). Androgen treatment effects on motor function, cognition, and behavior in boys with klinefelter syndrome4. *J. Pediatr.* 185, 193.e–199.e. doi: 10.1016/j.jpeds.2017.02.036
- Ross, J. L., Roeltgen, D. P., Stefanatos, G., Benecke, R., Zeger, M. P. D., Kushner, H., et al. (2008). Cognitive and motor development during childhood in boys with Klinefelter syndrome. *Am. J. Med. Genet. A.* 146A, 708–719. doi: 10.1002/ajmg.a.32232
- Ross, J. L., Stefanatos, G., Roeltgen, D., Kushner, H., and Cutler, G. B. (1995). Ullrich-Turner syndrome: neurodevelopmental changes from childhood through adolescence. *Am. J. Med. Genet.* 58, 74–82. doi: 10.1002/ajmg.1320580115
- Ross, J. L., Stefanatos, G. A., Kushner, H., Bondy, C., Nelson, L., Zinn, A., et al. (2004). The effect of genetic differences and ovarian failure: intact cognitive function in adult women with premature ovarian failure versus turner syndrome. *J. Clin. Endocrinol. Metab.* 89, 1817–1822. doi: 10.1210/jc.2003-031463
- Ross, J. L., Stefanatos, G. A., Kushner, H., Zinn, A., Bondy, C., and Roeltgen, D. (2002). Persistent cognitive deficits in adult women with Turner syndrome. *Neurology* 58, 218–225. doi: 10.1212/WNL.58.2.218
- Rovet, J. (2004). Turner syndrome: genetic and hormonal factors contributing to a specific learning disability profile. *Learn. Disabil. Res. Pract.* 19, 133–145. doi: 10.1111/j.1540-5826.2004.00097.x
- Rovet, J., Netley, C., Keenan, M., Bailey, J., and Stewart, D. (1996). The Psychoeducational profile of boys with klinefelter syndrome. *J. Learn. Disabil.* 29, 180–196. doi: 10.1177/002221949602900208

- Russell, H. F., Wallis, D., Mazzocco, M. M. M., Moshang, T., Zackai, E., Zinn, A. R., et al. (2006). Increased prevalence of ADHD in Turner syndrome with no evidence of imprinting effects. *J. Pediatr. Psychol.* 31, 945–955. doi: 10.1093/jpepsy/ksj106
- Skakkebaek, A., Gravholt, C. H., Rasmussen, P. M., Bojesen, A., Jensen, J. S., Fedder, J., et al. (2013). Neuroanatomical correlates of Klinefelter syndrome studied in relation to the neuropsychological profile. *NeuroImage Clin.* 4, 1–9. doi: 10.1016/j.nicl.2013.10.013
- Skakkebaek, A., Moore, P. J., Pedersen, A. D., Bojesen, A., Kristensen, M. K., Fedder, J., et al. (2017). The role of genes, intelligence, personality, and social engagement in cognitive performance in Klinefelter syndrome. *Brain Behav.* 7:e00645. doi: 10.1002/brb3.645
- Stergiakouli, E., Langley, K., Williams, H., Walters, J., Williams, N. M., Suren, S., et al. (2011). Steroid sulfatase is a potential modifier of cognition in attention deficit hyperactivity disorder. *Genes Brain Behav.* 10, 334–344. doi: 10.1111/j.1601-183X.2010.00672.x
- Tamm, L., Menon, V., and Reiss, A. L. (2003). Abnormal prefrontal cortex function during response inhibition in turner syndrome: functional magnetic resonance imaging evidence. *Biol. Psychiatry* 53, 107–111. doi: 10.1016/S0006-3223(02)01488-9
- Tartaglia, N. R., Wilson, R., Miller, J. S., Rafalko, J., Cordeiro, L., Davis, S., et al. (2017). Autism spectrum disorder in males with sex chromosome aneuploidy: XXY/Klinefelter Syndrome, XYY, and XXYY. *J. Dev. Behav. Pediatr.* 38, 197–207. doi: 10.1097/DBP.0000000000000429
- van Rijn, S., Bierman, M., Bruining, H., and Swaab, H. (2012a). Vulnerability for autism traits in boys and men with an extra X chromosome (47,XXY): the mediating role of cognitive flexibility. *J. Psychiatr. Res.* 46, 1300–1306. doi: 10.1016/j.jpsychires.2012.06.004
- van Rijn, S., Swaab, H., Baas, D., de Haan, E., Kahn, R. S., and Aleman, A. (2012b). Neural systems for social cognition in Klinefelter syndrome (47,XXY): evidence from fMRI. *Soc. Cogn. Affect. Neurosci.* 7, 689–697. doi: 10.1093/scan/nsr041
- van Rijn, S., Stockmann, L., Borghgraef, M., Bruining, H., van Ravenswaaij-Arts, C., Govaerts, L., et al. (2014). The social behavioral phenotype in boys and girls with an extra x chromosome (klinefelter syndrome and trisomy x): a comparison with autism spectrum disorder. *J. Autism Dev. Disord.* 44, 310–320. doi: 10.1007/s10803-013-1860-5
- Wistuba, J. (2010). Animal models for Klinefelter's syndrome and their relevance for the clinic. *Mol. Hum. Reprod.* 16, 375–385. doi: 10.1093/molehr/gaq024
- Wistuba, J., Luetjens, C. M., Stukenborg, J.-B., Poplinski, A., Werler, S., Dittmann, M., et al. (2010). Male 41, XXY* mice as a model for klinefelter syndrome: hyperactivation of leydig cells. *Endocrinology* 151, 2898–2910. doi: 10.1210/en.2009-1396
- Zhao, C., and Gong, G. (2017). Mapping the effect of the X chromosome on the human brain: neuroimaging evidence from Turner syndrome. *Neurosci. Biobehav. Rev.* 80, 263–275. doi: 10.1016/j.neubiorev.2017.05.023

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Aarde, Hrnčir, Arnold and Jentsch. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.