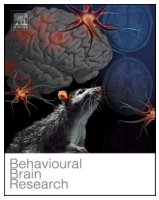
Behavioural Brain Research 461 (2024) 114860

Contents lists available at ScienceDirect

Behavioural Brain Research

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



Research report

Track-by-Day: A standardized approach to estrous cycle monitoring in biobehavioral research

Gianna M. Raimondi [a,b], Ashley K. Eng [a], Murphy P. Kenny [a], Madison A. Britting [a], Linnaea E. Ostroff [a,b,c,*]

[a] Department of Physiology and Neurobiology, University of Connecticut, Storrs, CT, USA
[b] Connecticut Institute for the Brain and Cognitive Science, University of Connecticut, Storrs, CT, USA
[c] Institute of Materials Science, University of Connecticut, Storrs, CT, USA

ARTICLE INFO

Keywords:
Estrous cycle
Sex differences
Fear conditioning

ABSTRACT

Despite known sex differences in brain function, female subjects are underrepresented in preclinical neuroscience research. This is driven in part by concerns about variability arising from estrous cycle-related hormone fluctuations, especially in fear- and anxiety-related research where there are conflicting reports as to whether and how the cycle influences behavior. The inconsistency may arise from a lack of common standards for tracking and reporting the cycle as opposed to inherent unpredictability in the cycle itself. The rat estrous cycle is conventionally tracked by assigning vaginal cytology smears to one of four qualitatively-defined stages. Although the cytology stages are of unequal length, the stage names are often, but not always, used to refer to the four cycle days. Subjective staging criteria and inconsistent use of terminology are not necessarily a problem in research on the cycle itself, but can lead to irreproducibility in neuroscience studies that treat the stages as independent grouping factors. We propose the explicit use of cycle days as independent variables, which we term Track-by-Day to differentiate it from traditional stage-based tracking, and that days be indexed to the only cytology feature that is a direct and rapid consequence of a hormonal event: a cornified cell layer formed in response to the pre-ovulatory 17β-estradiol peak. Here we demonstrate that cycle length is robustly regular with this method, and that the method outperforms traditional staging in detecting estrous cycle effects on Pavlovian fear conditioning and on a separate proxy for hormonal changes, uterine histology.

1. Introduction

Despite well-documented sex differences in many aspects of brain structure and function [18,55], female subjects have historically been excluded from routine preclinical research in the biobehavioral sciences. Research on rats, a preferred model for behavioral studies, is especially biased towards males; in 2017, 53.5% of studies in high-profile neuroscience journals and 82.2% of studies in behavior-focused journals used only male rats [35,43]. Anxiety-related disorders are more common in women than men [3,62], so studying anxiety-related behaviors in female subjects should provide better insight into brain mechanisms as well as greater translational relevance. Exclusion of females is often justified on the grounds that hormone fluctuations across the estrous cycle will increase variability in experimental data [60], and although this concern is controversial in the broader neuroscience literature [8,9,57,67,78], studies of sex and estrous effects on anxiety-related behaviors do report inconsistent results [22,41,58]. Disagreement between studies can fuel the perception that the estrous cycle is an uncontrollable confound that leads to irreproducible results, but it may simply be driven by the lack of universal standards for tracking and reporting the cycle [41]. Conventional methods of cycle tracking in rodents are complex and require specialized expertise [20,61,73]. Moreover, they are not tailored to the specific concerns of behavioral neuroscience studies. A streamlined, common framework for estrous tracking in behavioral research would improve reproducibility between studies and eliminate an unnecessary barrier to the widespread study of female subjects.

Sex hormones are known to affect brain function [29], and concerns about the estrous cycle in behavioral research relate mainly to two steroid hormones, 17β-estradiol and progesterone, whose serum levels change several-fold over the course of the cycle [15,51,70]. Direct measurement of hormones requires blood collection, which is invasive

* Corresponding author at: Department of Physiology and Neurobiology, University of Connecticut, Storrs, CT, USA.
E-mail address: linnaea.ostroff@uconn.edu (L.E. Ostroff).

https://doi.org/10.1016/j.bbr.2024.114860
Received 29 September 2023; Received in revised form 3 January 2024; Accepted 6 January 2024
Available online 10 January 2024

and stressful, so cycle progression is usually monitored by observing changes in vaginal cytology. Traditional tracking methods assign vaginal smears to one of four classic stages of mammalian reproduction: estrus, metestrus, diestrus, and proestrus [25,33,40,71]. The stages are defined qualitatively and do not have discrete boundaries, so interpreting smears is a subjective art that requires training and experience [7,20,44,61,73,76]. Subjective staging is not necessarily a problem; for example, the estrous cycle is often tracked in toxicology studies to determine whether it is disrupted by pollutants. In this context tracking methods need only be consistent within a study, and a popular guide's advice that "individual laboratories define their respective processes" for staging is appropriate and practical [20]. In contrast, the estrous cycle is not a dependent variable in behavioral studies. Because circulating steroid hormones may affect a given behavior, estrous stages are used as experimental grouping factors and it is easy to see how studies that define stages differently could report different results.

Semi-quantitative approaches [16,34,54] and machine learning tools [63,75] have been developed to reduce subjectivity in cytology interpretation. These tools use different terminology and criteria for the stages, however, so only studies using the same tool can be compared. In addition, reporting practices are not standardized. The rat estrous cycle typically lasts four days, and while the estrus stage lasts about 24 h, the other three vary in length from less than twelve hours to more than two days. The four stage names are sometimes used to refer to the four days of the cycle, and although some studies are explicit about this convenient use of terminology (e.g. [11,12]) it is more often ambiguous whether subjects are grouped by stages or days. It is also common practice to exclude subjects with "irregular" cycles, but there are no agreed-upon criteria for regularly versus irregularly cycling subjects and the number of excluded subjects is usually not reported. The absence of shared definitions leads to inconclusive data as to whether and when the estrous cycle is a source of variability [41] as well as the discouraging idea that most cycles are irregular [61]. These are arguably reasonable grounds for excluding female subjects, underscoring the urgent need for a widely accessible, streamlined tracking framework.

The implicit rationale for cytology staging in behavioral research is that the stages reflect serum levels of neurobiologically-relevant hormones, but the stages were defined long before the hormone cycle was understood. Only one change in the vaginal epithelium is known to be directly and temporally linked to a hormonal change: bulk desquamation of the stratum corneum the day after the preovulatory surge in 17β-estradiol [4,14]. This event defines one of the cytology stages, estrus. Estrus cytology is much more easily recognized than the other traditional cytology stages, as evidenced by the high concordance between experimenters and the performance of automated classifiers in detecting this stage [63,75]. Estrus cytology occurs on one day of each four-day cycle, so assigning cycle days relative to the day of estrus would be a simple alternative to daily staging by nuanced, subjective reading of cytology smears. Indeed, this type of strategy is used to monitor the cycle of golden hamsters, a popular model in reproductive biology whose cycle days are counted from a mucous discharge that occurs on the day of ovulation in that species [38,47,52,56]. Here we describe the use of an indexed cycle timing method, which we term Track-by-Day to distinguish it from conventional stage-centric tracking, in the rat. We evaluated the predictability of estrus cytology by quantifying vaginal smears under multiple conditions, and compared Track-by-Day with traditional staging for detection of estrous effects on Pavlovian fear conditioning behavior and correlation with hormone-sensitive features of uterine histology. The simplicity of Track-by-Day will make estrous cycle monitoring available to a greater number of researchers, and its use alongside or instead of traditional staging will allow comparisons between studies and metanalyses that should accelerate our understanding of the female brain.

2. Materials and Methods

2.1. Subjects

Subjects were adult female and male Sprague-Dawley rats (Hilltop Lab Animals Inc., Scottsdale, PA) pair-housed with ad libitum food and water. After arrival from the vendor, animals were acclimated to the housing room for a minimum of five days before experiments began, or a minimum of ten days for those housed with a reverse light cycle. Housing rooms were lit with a dim red lamp (3 W, 120 V LED) to allow personnel to work during the dark period without exposing rats to white light during the designated dark period. All animal procedures were approved by the Institutional Animal Care and Use Committee at the University of Connecticut.

Experiment 1: Twelve adult female Sprague Dawley rats were habituated to the colony room for one week before handling or swabbing began at PND 92. The colony room was on a 14:10 light cycle (on 07:00, off 21:00); other females, males, and breeding pairs were also present in the colony. Females were swabbed once a day (08:00–10:00) for two cycles. Estrous stages were assigned by subjectively interpreting vaginal cytology stained with H&E as per Becker et al. [7]. Later, cytology was quantified using Reconstruct software by at least two blinded experimenters.

Experiment 2: Sixteen adult female Sprague Dawley rats (PND ~35) were habituated to the colony room for five days before any handling or swabbing. The colony room was on a 12:12 light cycle (on 07:00, off 19:00); other females, some pregnant or nursing, were also present in the colony. Females were swabbed twice a day (08:50–09:50 or 11:00–13:00, and 14:00–17:30) for five cycles. Estrous stages were assigned by subjectively interpreting vaginal cytology stained with H&E, as per Becker et al. [7].

Experiment 3: Six adult female Sprague Dawley rats (PND 45) were habituated to the colony room for one week before any handling or swabbing. The colony room was on a 12:12 reverse light cycle (off 09:00, on 21:00); other females, males, and breeding pairs were also present in the colony. Females were swabbed three times a day, twice in the dark cycle under dim red light (09:30–12:30 and 16:30–18:30) and once in the light cycle (22:30–23:30) for five cycles. Estrous stages were assigned by quantifying vaginal cytology stained with Shorr stain (Sigma) and modified from Paccola et al. (2013).

Experiment 4: Sixty adult female and twenty-four male Sprague Dawley rats (PND 30) were habituated to the colony room for two weeks before any handling or swabbing. The colony room was on a 12:12 reverse light cycle (off 09:00, on 21:00); other females, males and breeding pairs were also present in the colony. Females were swabbed once per day in the dark cycle (12:00–14:00) under dim red light.

2.2. Collection of vaginal smears

Animals were gently restrained and a saline-dipped cotton-tipped applicator (2 mm diameter, cotton tip length 1 cm) was inserted into the vagina to the depth of the cotton tip. Swabs were gently rolled out onto a gelatin-subbed microscope slide and the smears were allowed to dry before any staining. Male rats were subjected to identical handling, and males and females were handled by different experimenters on the same day.

2.3. Staining, imaging, and analysis of vaginal smears

For cresyl violet staining, slides were dipped in 0.1% aqueous cresyl violet (Millipore Sigma, Inc.) for 5 min, then rinsed in water. H&E (Vector Labs, Inc.) and Shorr (Millipore Sigma, Inc.) stains were used according to the manufacturers' instructions. Smears used for quantification were stained with H&E except for the terminal smears taken after the behavior experiment, which were Shorr stained. Stained slides were dehydrated in ethanol, cleared in toluene, and coverslipped with DPX

mounting medium. Smears were imaged at 10X using either a Keyence BZ-X700 microscope or on a compound microscope with a Canon EOS 800D digital camera. Reconstruct software [27] was used for cell quantification. A counting frame of 285.75 mm × 285.75 mm was applied to a representative area of each image, and all cells within the counting frame were counted and classified. Three separate images were analyzed per smear by at least two blinded experimenters.

Impedance measurement.

To construct the probe, two gold pins (3 cm long by 1 mm wide) were soldered to copper wire with lead-free silver solder, and heat shrink was used to secure them in a parallel position 3 mm apart. For impedance readings, the probe was attached to a standard multi-meter and inserted to a depth of ~2 mm so that the pins were in contact with the dorsal and ventral vaginal wall. Impedance readings were taken prior to swabbing for cytology smears, and the probe was sanitized with 100% ethanol and allowed to dry between readings.

2.4. Pavlovian fear conditioning

All behavioral experiments were conducted in the dark period, 3–7 h after lights-off. The behavioral apparatus, training, and testing protocols were identical to those used in Ostroff et al. [53] with the exception that the conditioning chambers were unlit during all procedures. Briefly, animals were habituated to the conditioning chamber (Coulbourn Instruments) for 30 min on each of two consecutive days before training. Tones (30 s, 5 kHz, 80 dB) were delivered through a speaker in the chamber and shocks (1 s, 0.7 mA) were delivered through a grid floor. Paired training consisted of five tones co-terminating in footshocks over a single 32.5-minute training session, and unpaired training consisted of five non-overlapping tones and shocks. Testing consisted of three tones presented in the same chamber, with the context modified by the addition of a smooth opaque acrylic floor and the scent of peppermint. Cage-mate pairs were trained and tested at the same time using the same protocol, and because pairs of females did not synchronize their cycles it was impossible to collect equal numbers of each stage. Freezing during the tones (CS), and 30 s prior to tone (Pre-CS), was scored by three experimenters blinded to condition, sex, and stage, and the scores were averaged for analysis. Darting was also scored by blinded experimenters using criteria per [32].

2.5. Shock sensitivity

Shock sensitivity was assessed in a separate group of subjects that were not used in other experiments. Adult (PND 90) female (n = 26) and male (n = 8) rats were habituated to the reverse light-cycle housing room for at least two weeks. Females were tracked for four cycles leading up to the experiment and testing was divided across four days to ensure that all cycle days were represented. Males were match-handled and run separately after the females. Shock sensitivity was performed with the same equipment used for Pavlovian conditioning during the mid-dark cycle. Rats were placed singly in a testing box and a series of ten shocks were presented at 40 s intervals. Shock intensity began at 0.1 mA and increased in steps of 0.1 mA, so that the final shock was 1 mA. Cage-mate pairs were tested in parallel. The intensity of the first shock evoking a shuffle was recorded for each animal. Shuffles were defined as a fast, frenzied movement; some animals may shuffle in place by quickly moving all four paws.

2.6. Collection of uterine tissue

One hour after long-term memory testing, rats were deeply anesthetized with chloral hydrate (750 mg/kg i.p.) and perfused transcardially with 4% paraformaldehyde and 0.1% glutaraldehyde in 0.1 M phosphate buffer (pH 7.4) at room temperature. The uterine horns were removed and post-fixed overnight in 4% paraformaldehyde, then stored in phosphate buffered saline with 0.01% sodium azide until further processing.

Segments 0.5 to 1 cm long and roughly 1.5 to 2 cm away from the oviduct were dissected for embedding. Samples were dehydrated in ascending concentrations of ethanol, then infiltrated with a 1:5 mixture of methyl methacrylate and butyl methacrylate with 0.5% benzoin methyl ether. Samples were cured in gelatin capsules under UV light for 48 h in a freeze-substitution unit (Leica Microsystems). All dehydration, embedding, and curing were performed at 0ºC. The cured blocks were trimmed to expose a full cross section of the uterus, and 1 μm sections were cut on an ultramicrotome (Leica Microsystems). Sections were collected on gelatin-subbed slides, stained with 0.5% toluidine blue in 1% sodium borate for five minutes, and coverslipped with DPX. One representative cross-section was imaged per rat at 10X on a Keyence BZ-X700 microscope, and additional 40X images were taken on a compound microscope with a Canon EOS 800D digital camera.

2.7. Analysis of uterine histology

Uterine area, endometrium area, myometrium area, luminal epithelium area, lumen dilation, and number and size of endometrial glands were measured on 10X images using Reconstruct software. Mitotic figures and tissue necrosis were quantified in 40X images. Mitotic figures were counted in the luminal epithelium and the epithelium of the endometrial glands and normalized to area. Each analysis was performed by at least two blinded experimenters.

2.8. Statistics

Means of more than two groups were compared using one-way ANOVAs, and effects significant at $p < 0.05$ were followed with a Bonferroni post-hoc test. Unpaired t-tests were used when comparing two groups. Full results of statistical comparisons are given in Supplemental Table 1.

3. Results

3.1. Detection of directly hormone-driven changes in vaginal cytology

An estrous tracking strategy that infers cycle progression from an index point requires a stereotyped hormone cycle with an unambiguously detectable, temporally discrete proxy for a single cycle event. Multiple quantitative studies have reported a consistent pattern of circulating sex hormones across the rat estrous cycle [15,51,70]. Both 17β-estradiol and progesterone are lowest on the day of ovulation and both undergo a rapid rise and fall on the fourth day, with progesterone peaking a few hours after 17β-estradiol (Fig. 1a). Both hormones have direct effects on the vaginal epithelium (Fig. 1b). The preovulatory surge in 17β-estradiol stimulates cell proliferation and keratinization, creating a cornified cell layer (stratum corneum) that may be accompanied by leukocyte infiltration into the upper layers of the epithelium. Keratinization ceases with the fall in 17β-estradiol, and the ensuing progesterone peak induces cell proliferation with mucosal differentiation, after which the epithelium enters an atrophic state until 17β-estradiol rises again [4,25,31,39,49].

Changes in the epithelium are visible in smears, and conventional staging is generally based on proportions of three cell types: cornified cells (Fig. 1c), which are flattened, enucleated, and fully keratinized; nucleated (non-keratinized) epithelial cells (Fig. 1d); and leukocytes (Fig. 1e), which are small and have multi-lobulated nuclei. Bulk delamination of the stratum corneum occurs within 24 h of the 17β-estradiol surge [15,51,70], and smears dominated by cornified cells define the estrus stage. The ensuing stages vary in length (Fig. 1f) and are defined by rough proportions of lingering cornified cells, leukocytes, and nucleated epithelial cells [7,16,20,44,73,76]. Formation of the stratum corneum is a direct and immediate consequence of the pre-ovulatory 17β-estradiol peak, and delamination of cornified cells should thus be a reliable indicator of cycle timing. Our first goal was therefore to ensure that it is unmistakable in smears.

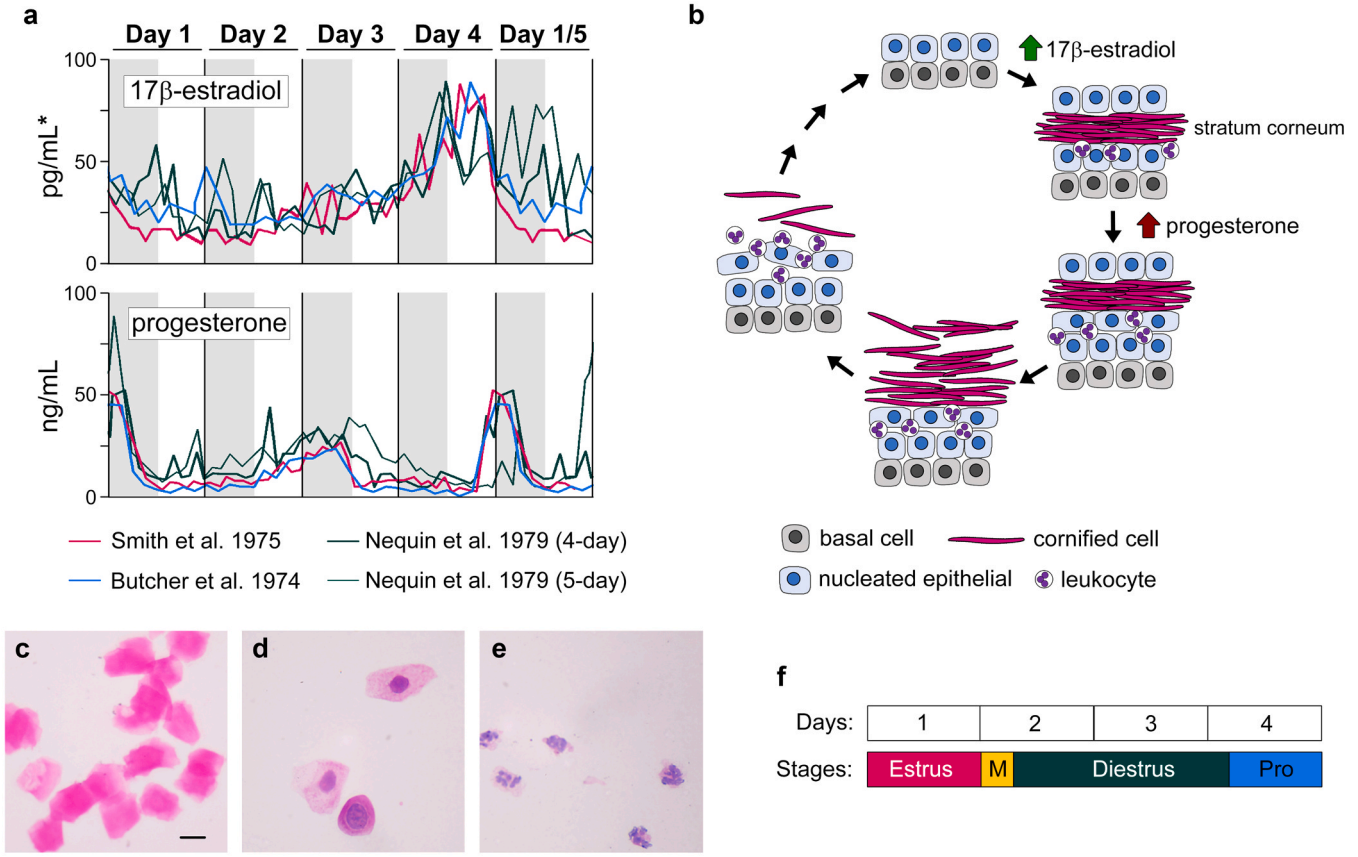


Fig. 1. Estrous cycle progression. (**a**) Serum concentrations of 17β-estradiol and progesterone over four- and five-day cycles in the rat as reported in three studies. Gray regions indicate the 12-hour dark period of each day, except for the Nequin et al. study which used a 10-hour dark period. Day 1 is the day of ovulation and termed the estrus day in all studies. *Axis values from Ref. 21, with data from Refs. 19 and 20 plotted at 66% and 200% respectively. (**b**) Schematic of changes in the vaginal epithelium in response to the brief surges of serum hormones during the estrous cycle. (**c-e**) Examples of the three cell types in a vaginal smear stained with H&E: (c) cornified, (d) nucleated epithelial, and (e) leukocytes. Scale: 10 μm. (**f**) The four terms that denote the progression of the rat estrous cycle may be used to refer to days (top row) or to cytology stages (bottom row).
Data adapted from Butcher et al. [15], Smith et al. [70], and Nequin et al. [51].

Histological stains provide contrast between cell types on vaginal smears and are often used to facilitate estrous staging [7,34,68], so we evaluated three stains – hematoxylin and eosin (H&E), Shorr stain, and cresyl violet – for their ability to distinguish cornified cells from other types. Even without staining, smears containing clumps of delaminated cornified cells (Fig. 2a) could be distinguished from smears containing a more disperse mix of cell types (Fig. 2e). H&E staining allowed rapid, unambiguous differentiation of smears consisting entirely of cornified cells (Fig. 2b) from smears containing a mix of cell types (Fig. 2f) due to the contrast between the bright pink cornified cells (Fig. 1c) and the purple nuclear staining of the other cells (Fig. 1d-e). Shorr stain also readily revealed contrast between cornified cells (Fig. 2c) and other cell types (Fig. 2g). Although cornified cell clumps were visible with cresyl violet staining (Fig. 2d), it did not produce chromatic contrast between cell types (Fig. 2h), and thus offered only a marginal improvement versus unstained smears. We chose H&E as our default stain because it produces clear contrast and is also more commonly used than Shorr stain.

3.2. *Experiment 1: Cornified cells in vaginal smears peak regularly across subjects and cycles*

To determine whether cornified cell desquamation could be used to time the cycle, we collected daily vaginal smears from twelve female rats for eight days and quantified the proportion of each cell type. Cornified cells were present in every smear and in some were the only cell type present, consistent with bulk desquamation of the stratum corneum (Fig. 2a-d). Within the first three days, a single read of 100% cornified cells (a cornified peak) was seen in all twelve rats (Fig. 3a). The cornified peak did not occur on the same calendar day for all animals, but when reads were aligned by the first peak a second occurred four days later in eleven of the twelve subjects (Fig. 3b). Cornified peaks occurred simultaneously in only two of the six cage-mate pairs, consistent with reports that rats do not synchronize cycles [65]. We designated the day of a cornified peak as Day 1 and numbered the other days of the cycle accordingly. Eight consecutive days of swabbing ensured each cycle day occurred twice. When the percentage of cornified cells on each read was averaged by cycle day, it was higher on Day 4 than on Days 2 and 3, but variability was high on all three days (Fig. 3c). To examine within-rat variability across cycles, the absolute value of the difference between the two cycles was calculated for each cycle day. Day 1 was the most consistent between cycles, as expected, and Day 4 the least (Fig. 3d).

Nucleated epithelial cells were present in most smears on Days 2–4 (Fig. 3e) but showed high variability when aligned by the cornified peak (Fig. 3f) with the highest percentage on Day 3 (Fig. 3g). Within-rat variability was similarly high after Day 1 (Fig. 3h). Leukocytes were detected less often than the other cell types (Fig. 3i). Some cycles did have a surge in leukocytes on Day 2 (Fig. 3j), consistent with the traditional metestrus stage (Fig. 1f), which is defined by copious leukocytes [20,73]. A leukocyte peak did not occur in all rats (Fig. 3k) or consistently within rats across the two cycles (Fig. 3l). Vaginal impedance has been reported to change over the estrous cycle and has been

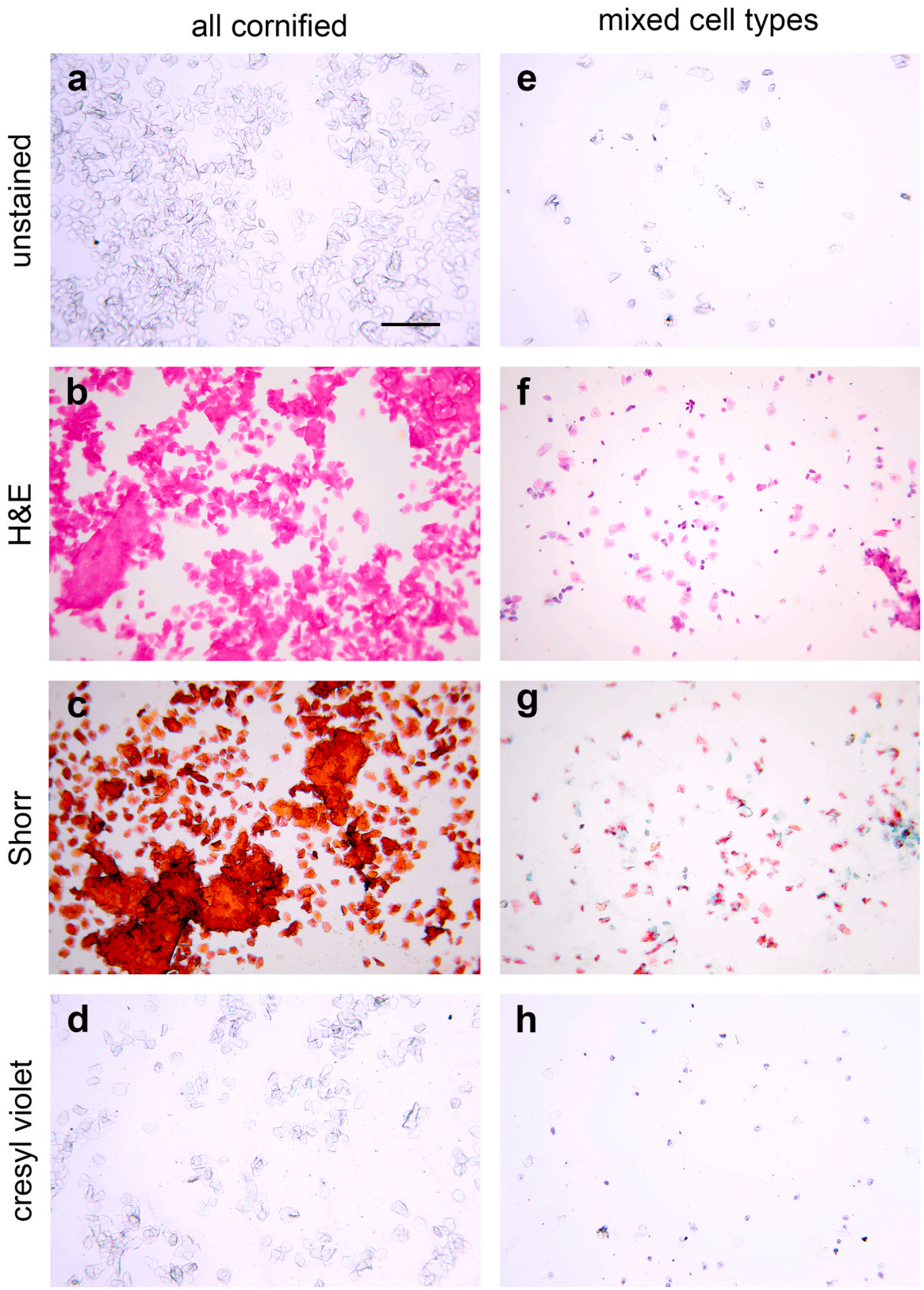


Fig. 2. Vaginal cytology under different staining conditions. (a-d) Smears consisting only of cornified cells. (e-h) Smears containing a mix of cell types. Contrast between cell types is low in unstained smears (a,e), high in smears stained with hematoxylin & eosin (H&E; b,f) or Shorr stain (c,g), and low in smears stained with cresyl violet (d,h). Scale: 100 μm.

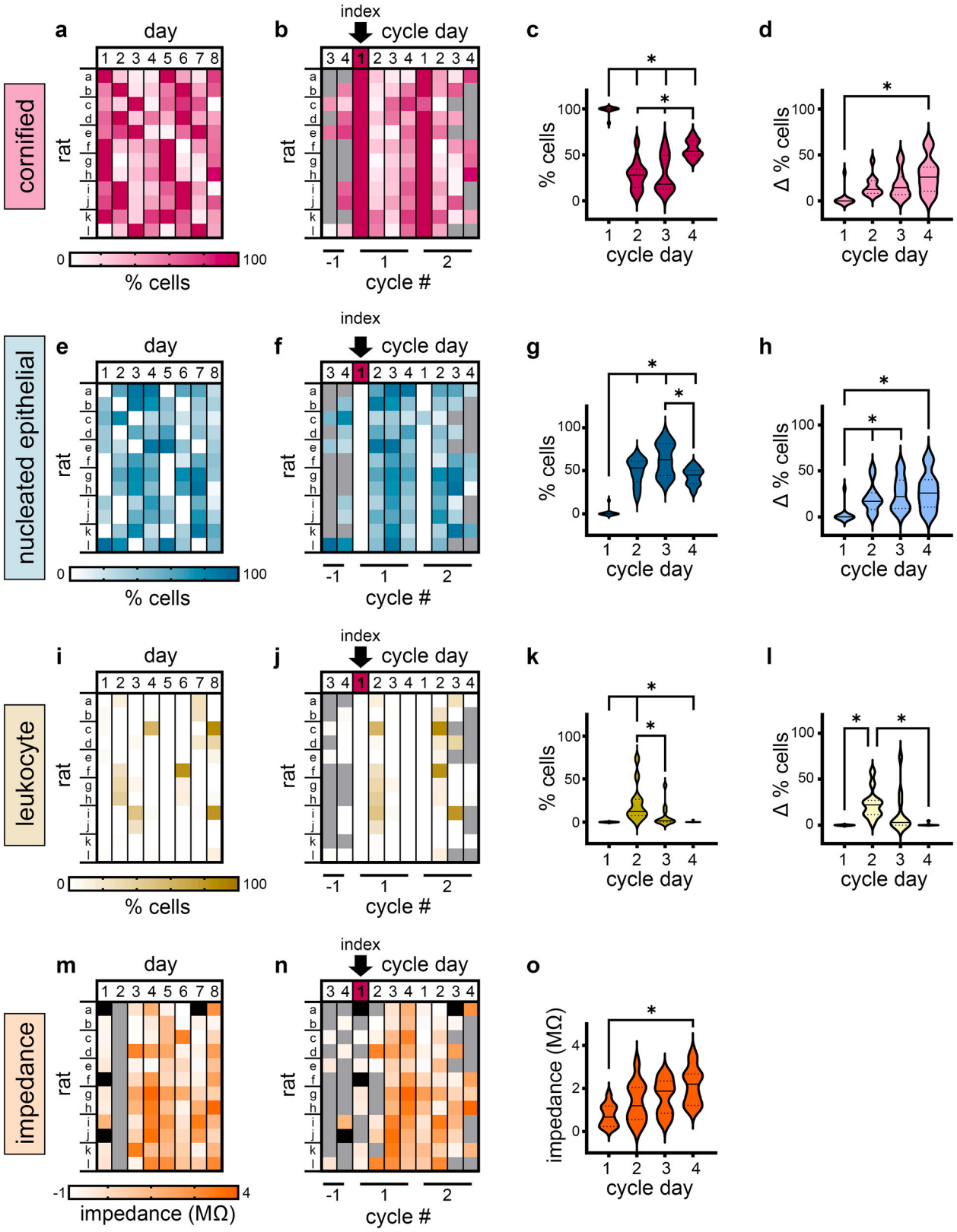


Fig. 3. Vaginal cell types across days. (a) Cornified cells as a percentage of all cells for twelve rats over eight days. Gradient scale also applies to (b). (b) Progression of cornified cell percentage aligned by each rat's first peak (100%). (c) Percent cornified cells by cycle day, averaged across rats according to cycle day. (d) Absolute value of the difference in % cornified cells between the two cycles. (e-l) Data for nucleated epithelials (e-h) and leukocytes (i-l) in the same format as data in (a-f). (m) Vaginal impedance (MΩ) collected on the same eight days. Impedance readings were not taken for the second day of tracking (gray), and a few individual readings were inconclusive (black). Gradient scale also applies to (n). (n) Impedance measures aligned by the day of the first cornified peak as in (b). (o) Mean impedance on each day of the second cycle only. For all raster plots, lines on y-axis indicate cage-mate pairs. * $p < 0.05$.

advocated as an easy alternative to cytology for cycle tracking [5,17,37,59], although one study found no correlation between impedance and serum hormone levels [69]. We measured impedance at the time of most swab collections and found that it varied over the cycle in all animals (Fig. 3m). No clear cyclic pattern relative to the cornified cell peak was observed (Fig. 3n), although on average impedance was lower on Day 1 than Day 4 (Fig. 3o). Aligning the data to the first peak of either nucleated epithelial cells or leukocytes instead of cornified cells confirmed that only a few rats had a second peak of the same cell type four days later (Supplemental Figure 3–1).

3.3. Experiment 2: The cornified cell peak is regular across environmental conditions

The animals in our first experiment were housed with a 14-hour light period, which is common in breeding colonies but not typically used for behavior experiments. To ensure that the cornified peak is reliable in animals housed on a 12-hour light cycle, we began with a new cohort of 16 rats (PND 35). These animals were housed in non-ventilated cages in a room with only female rats and nursing litters. To determine whether cytology would be more predictable early or late in the day, swabs were collected twice per day for eight days, at two hours and eight hours after lights-on (L2 and L8). For the next eight days, swabs continued to be taken from 11 of the 16 rats. Impedance was measured and blood samples were collected from the lateral tail vein alongside many of the swabs in the second round. For the last four days, swabs were collected from all 16 rats once per day, six hours into the light period (L6). The proportion of each cell type was quantified on every smear (Supplemental Figure 4–1). During twice-daily collection, all 16 rats had regular, discrete cornified peaks consisting of 1 – 4 consecutive reads. Most four-day cycles had either one peak read at the L2 time point or two peak reads, usually L2 followed by L8. This pattern indicates that the morning (L2) read is more reliable than the afternoon (L8) read for tracking the cycle, and suggests that the cornified peak lasts about 24 h and begins sometime during the dark period. Because some rats had a cornified peak on the first day of sampling, we aligned the cornified cell data to the second peak to better visualize the distribution of surrounding reads (Fig. 4a). No patterns were evident in the other two cell types or in impedance (Supplemental Figs. 4–2), so only the cornified cells were used to track cycles. The intent of the blood collection was to verify levels of 17β-estradiol and progesterone, but tail vein sampling did not yield sufficient volumes of serum for either of two commercial ELISA kits we tested.

Most cycles lasted four days, but in contrast to the first experiment some five-day cycles were observed. Five-day cycles are generally characterized by two consecutive days of cornified peaks, consistent with the extended peak of 17β-estradiol reported for five-day cycles [51]. Clear five-day cycles were observed in five of the 16 rats, but all five also had four-day cycles (Fig. 4b). Four of the eleven rats that underwent blood draws shifted their cycle length by one day, and some did not have detectable peaks in the final round of smears that were collected at the L6 time point. This could reflect disrupted cycles, but it is also possible that peaks were missed if their timing was shifted away from mid-day. Peaks occurred on the same days for two of the eight cage-mate pairs, again demonstrating chance levels of synchrony within cages.

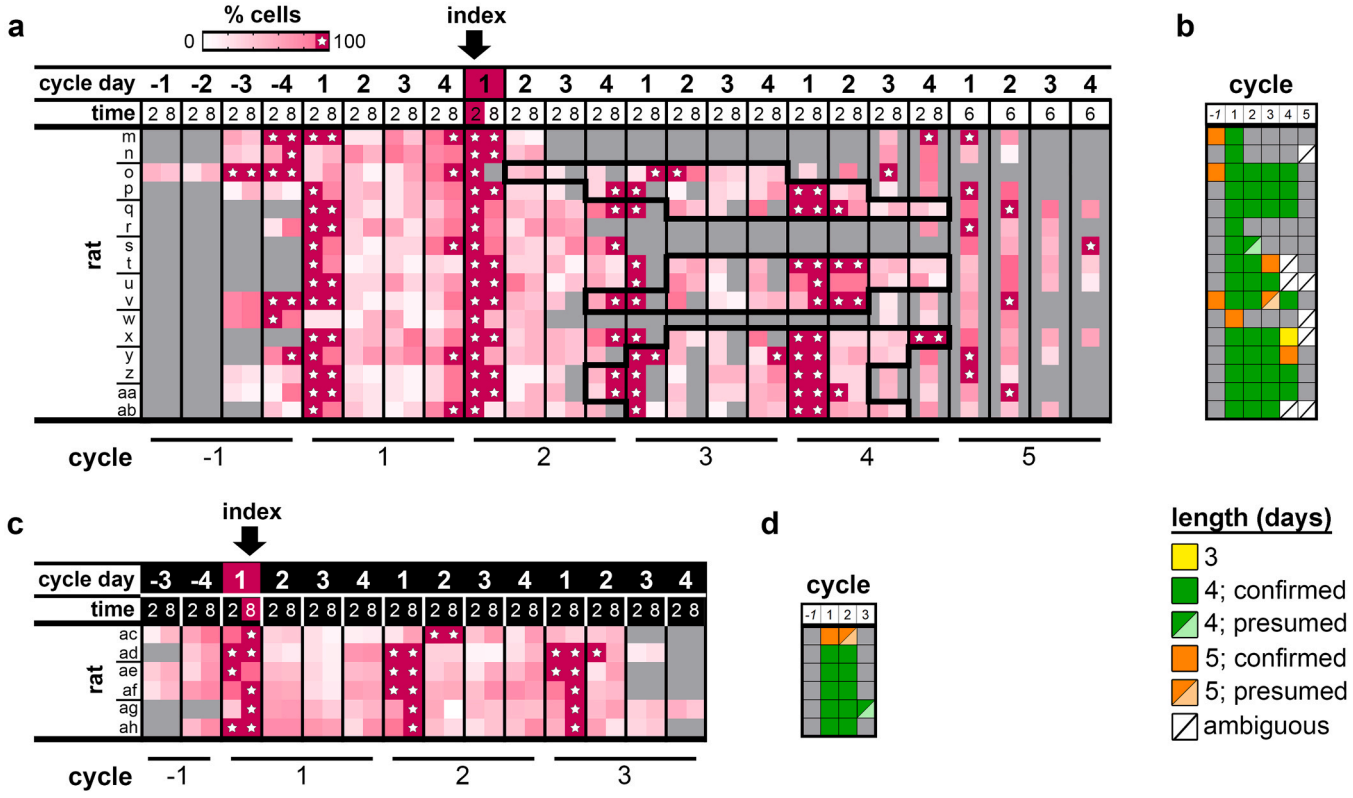


Fig. 4. Tracking cornified cells across the light-dark cycle. (a) Percent cornified cells in smears collected at time points 2, 6, or 8 h after lights-on from rats housed on a 12-hour light-dark cycle. Reads are indexed to the lights-on + 2 h peak read of each rat's second tracked cycle. Times when a smear was not collected are shown in gray. The black outline denotes days when tail vein blood draws were also performed on 11 of the 16 rats. (b) The majority of tracked cycles lasted four days, and every rat had at least one four-day cycle. A confirmed 4-day cycle was one in which all four days plus Days 1 and 2 of the next cycle were observed. A 4-day cycle was presumed if the following Day 2 smear was not available to exclude the possibility of a 5-day cycle. 4-day cycles were also presumed when the first or final few days of tracking were consistent with a partial 4-day cycle. (c) Percent cornified cells in smears collected twice per day, 2 and 8 h after lights-off, and indexed to the first peak 8-hour read. (d) Five of six animals had two consecutive four-day cycles, and the sixth had a five-day cycle followed by a second apparent five-day cycle. Lines on y-axis indicate cage-mate pairs.

3.4. Experiment 3: The cornified peak is reliably detected during the dark period

Rats are nocturnal, so behavior experiments performed in the dark period may be more ethologically relevant. To confirm that the cornified peak could be used to time the cycle during the dark period, we housed a new cohort of six rats (PND 41) in a room with a reverse 12-hour light cycle. This room contained a breeding colony and had ventilated cages, as in the first experiment. Smears were collected twice per day at two hours and eight hours after lights-off (D2 and D8) for twelve days, and cells were quantified for each smear (Supplemental Figure 4–3). Cornified peaks were detected more often in the D8 smears than in the D2 smears, so cycles were aligned to the first D8 peak read (Fig. 4c). When two consecutive peaks occurred, the first was more often a D2 read. This order plus the higher frequency of peak reads at D8 and L2 versus D2 and L8 indicates that the cornified peak commences around the light-to-dark transition. Only one of the six rats had a five-day cycle, versus eight four-day cycles amongst the other rats (Fig. 4d). Interestingly, two-day peaks were not observed in this group. Leukocytes were also very scarce relative to the smears collected during the light period, consistent with the reported short duration of the metestrus stage, which is defined by copious leukocytes.

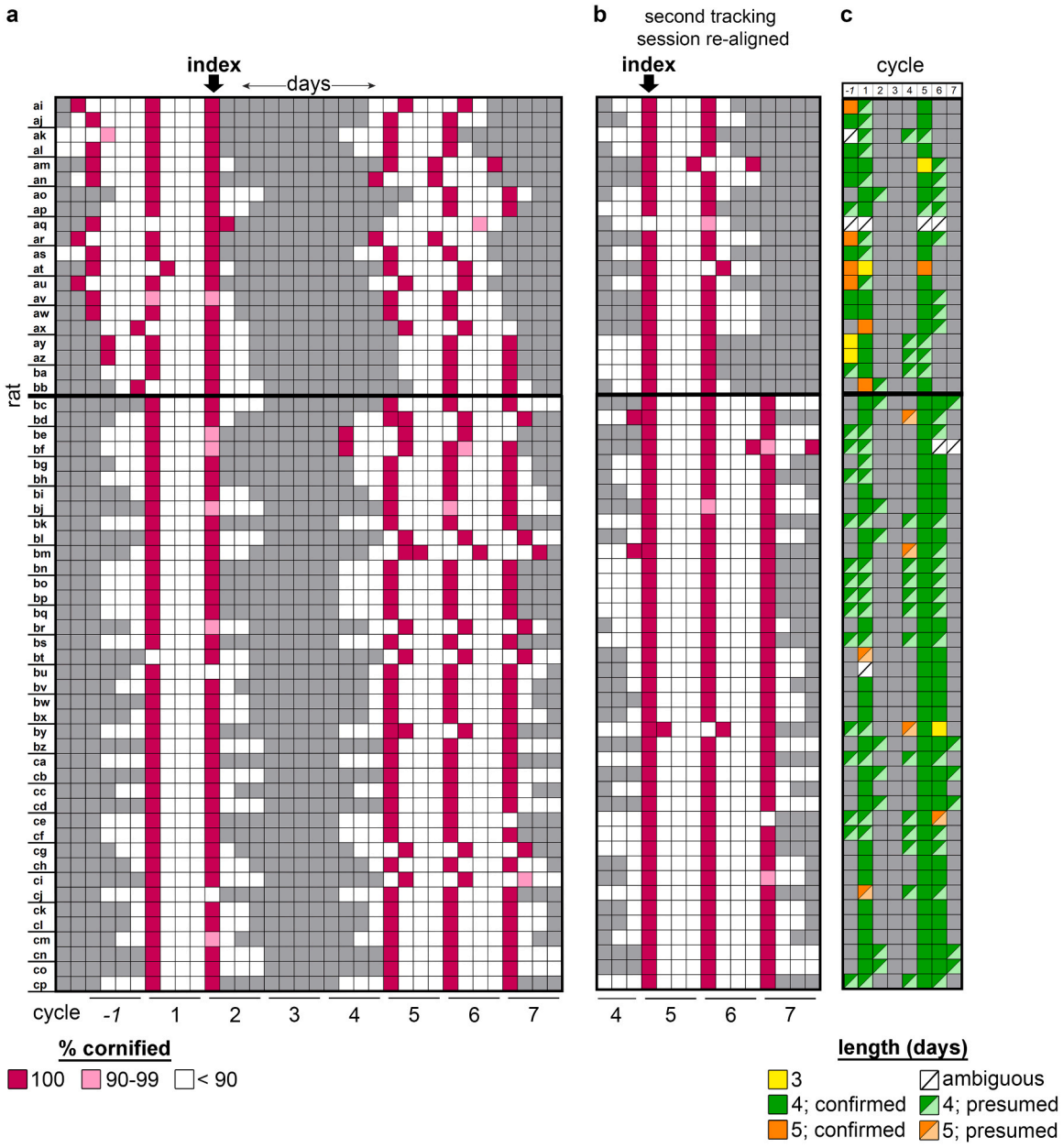


Fig. 5. Track-by-Day in a large cohort across four weeks. (a) Daily smears were collected over two rounds separated by a break. Two groups (n = 20 and n = 40) were tracked with slightly different collection and break schedules. Smears were categorized as either fully cornified or not, although in a few instances 90–99% cornification presented in the smear when a peak was expected (light pink). Cycles were indexed to the final cornified peak before the break. Gray boxes indicate days when no smear was collected, and lines on the y-axis show cage-mate pairs. (b) The second tracking round realigned using its first cornified peak as the index day. (c) Across 60 animals, 240 full cycles were tracked: the majority were four-day (211 of 237, 89.0%), while the rest were either three-day (5 of 237, 2.1%), five-day (13 of 237, 5.5%), or ambiguous (8 of 237, 3.4%). Importantly, by the second tracking session (PND 65), the incidence of four-day cycles increased (125 of 136, 91.9%), therefore 97% of females (58 of 60) were regularly cycling. Gray boxes in (c) refer to alignment resulting in an incomplete cycle, i.e., less than four or five days of cytology data.

3.5. Experiment 4: Days are more sensitive than stages in detecting estrous cycle effects on fear conditioning

We next compared the performance of the Track-by-Day method to traditional staging in a Pavlovian conditioning paradigm to compare fear learning across days of the estrous cycle, and between females and males. Smears were collected from females once per day, mid-dark period, for two rounds of 8–12 days, with a break of 8 or 10 days in between. Multiple tracking rounds allowed each subject's cycle to be observed over four weeks, and the break allowed us to interleave 60 female rats in a large experiment. Each smear was classified as either 100% cornified, signifying Day 1, or not. In a few cases, a read of 90–99% cornified cells was used to define Day 1 if it occurred on the expected day and no 100% cornified read was seen. When cycles were aligned by the first cornified peak they were mostly regular during the first round of swabbing. A few rats drifted during the break, almost always by one day (Fig. 5a). Realigning the cycles to the first cornified peak in the second tracking round showed that consecutive cycles are overwhelmingly regular (Fig. 5b). Over the entire experiment, the majority of cycles lasted four days (Fig. 5c). A total of 237 cycles were followed, of which 211 were four-day, 13 were five-day, and 5 were three-day.

The behavior experiment began 1 – 4 days after the final day of smear collection (Supplemental Figure 5–1). The final cycle was used to assign females pseudorandomly to a paired fear learning or unpaired control protocol to ensure representation of cycle days in each experimental group. A few females were assigned to a third tone-only group whose behavior data were not analyzed. Males were match-handled and balanced across experimental groups as well. Training and long-term memory testing were performed four days apart so that they would occur on the same day of a (four-day) cycle (Fig. 6a). We intentionally avoided training and testing on Day 1 for two reasons. First, our twice-daily swabbing experiments found that the cornified peak is most frequently observed in the late dark period and early light period (Fig. 4a, c), which would mean that our dark period Day 1 smears are taken during the pre-ovulatory drop in progesterone (Fig. 1a). This drop is rapid, and even with experiments performed within a 4-hour time window, there could be dramatic differences in hormone levels within a Day 1 experimental group. Second, five-day cycles are expected to have two consecutive days of high 17β-estradiol (Fig. 1a) and potentially two days of peak cornified reads (Fig. 4a).

In this case, a Day 1 smear and a Day 5 smear are indistinguishable without the following day's smear to confirm a four- or five-day cycle. A final smear was collected from each female after the long-term memory test, and although we attempted to predict and avoid Day 1, four final smears predicted to be Day 4 or Day 2 showed Day 1 cytology. These rats were excluded from behavior analysis.

During the long-term memory test, both males and females in the paired training group showed higher freezing to the tones relative to the pre-tone period. Females, but not males, in the unpaired group also showed higher freezing to the tone, indicating non-associative effects of training (Fig. 6b). Females on all cycle days showed increased freezing to the tone after paired training, but only unpaired subjects on Day 2 showed non-associative freezing (Fig. 6c). To compare Track-by-Day to traditional staging, a blinded, experienced experimenter used conventional cytology guidelines to classify the final smears for females that underwent behavior procedures as diestrus or proestrus. Stages markedly overlap, with Day 2 consisting mainly of diestrus reads and Days 3 and 4 of both diestrus and proestrus (Fig. 6d). As in the previous dark-period experiment, leukocytes were very scarce and thus no animals were assigned to the metestrus stage. When the female data were analyzed by stage, no difference was seen between diestrus and proestrus subjects (Fig. 6e). Relatively high freezing was seen in the unpaired group overall. Another cued fear conditioning study conducted during the dark period reported similar results [21], suggesting that context generalization may be higher in the dark period. To isolate tone responses, pre-tone freezing was subtracted from tone freezing. Both males and females had significantly higher freezing to the tone in the paired group (Fig. 6f), confirming associative learning to the tone. Among females, freezing in the paired and unpaired groups did not differ in subjects on Day 2 (Fig. 6g) or those classified as diestrus (Fig. 6h), consistent with the non-associative tone freezing in these unpaired groups. Overall freezing did not differ between males and females (Fig. 6i) but was significantly lower on Day 3 than on Days 2 and 4 (Fig. 6j). Diestrus and proestrus did not differ (Fig. 6k), meaning that Track-by-Day was more sensitive in detecting a cycle-related change in behavior.

Training was performed using the same shock parameters for all animals, but shock sensitivity has been reported to be affected by sex and hormonal state [6]. Because lower freezing on Day 3 and non-associative sensitization on Day 2 could be explained by lower and higher shock sensitivity, respectively, we evaluated shock sensitivity in a separate cohort of female and male rats. Females displayed higher sensitivity than males, but there were no differences between cycle days (Supplemental Figure 6–2).

3.6. Additional cell types and overall cell numbers associated with Day 2

The final smears from Experiment 4 were stained with Shorr stain, which reveals cytology in finer detail than H&E and allowed five cell types to be quantified instead of three. Cornified cells largely appear bright orange with Shorr stain, and the full cornification of Day 1 is easily detected (Fig. 7a). Nucleated epithelials show clear contrast between the nucleus and cytoplasm and high variability on Days 2–4 (Fig. 7b). Leukocytes were very rare in these smears (Fig. 7c), consistent with our earlier H&E-stained smears taken during the dark period (Supplemental Figure 4–3). Non-nucleated epithelial cells are difficult to differentiate with H&E staining but are very clear with Shorr stain. Acidophilic non-nucleated epithelials were very sparse (Fig. 7d). In contrast, basophilic non-nucleated epithelials, recognized by bright cyan staining in crisp round cells, appeared frequently on Days 3 and 4 but rarely on Day 2 (Fig. 7e). Independent of the specific cell types, the absolute number of cells per smear was higher on Day 2 than on Days 3 and 4, but there were no differences between stages (Supplemental Figure 7–1). Based on gradual changes in cytology, some smears were classified as late estrus (E_L), late diestrus (D_L), or late proestrus (P_L). Similar to the canonical stages (Fig. 6e), transitional stages progressed across days (Supplemental Figure 7–1). Number of cells per smear were not different when compared by stages of the estrous cycle (Supplemental Figure 7–2). Although not as unmistakable as the fully-cornified cytology of Day 1, these two features could be helpful in confirming Day 2 by large cell numbers or a paucity of basophilic non-nucleated epithelial cells.

3.7. Cycle days are more predictive of uterine histology than traditional stages

Unlike the vaginal epithelium, uterine tissue is sensitive to hormone changes throughout the estrous cycle [19,73], so if cycle days are a better proxy for hormone levels than cytology stages this should be reflected in uterine histology. To test this hypothesis, rats in the behavior experiment were perfused with fixative and the uterine tissue was harvested immediately after collection of the final post-long-term memory test smears. Gross examination of uterine sections across days (Fig. 8a) revealed a dramatic dilation of the lumen in the five rats with Day 1 cytology (Fig. 8b). The only other rat with this uterine morphology was the one that did not have any four- or five-day cycles and did not have a cornified peak during the final round of smear collection (Fig. 5c), a combination of persistent non-estrus cytology and histology that is consistent with pseudopregnancy.

Day 1 of the cycle can be easily identified by vaginal cytology (Fig. 2, Fig. 3) and uterine lumen dilation (Fig. 8). Quantification of uterine histology features revealed less dramatic, yet still detectable changes

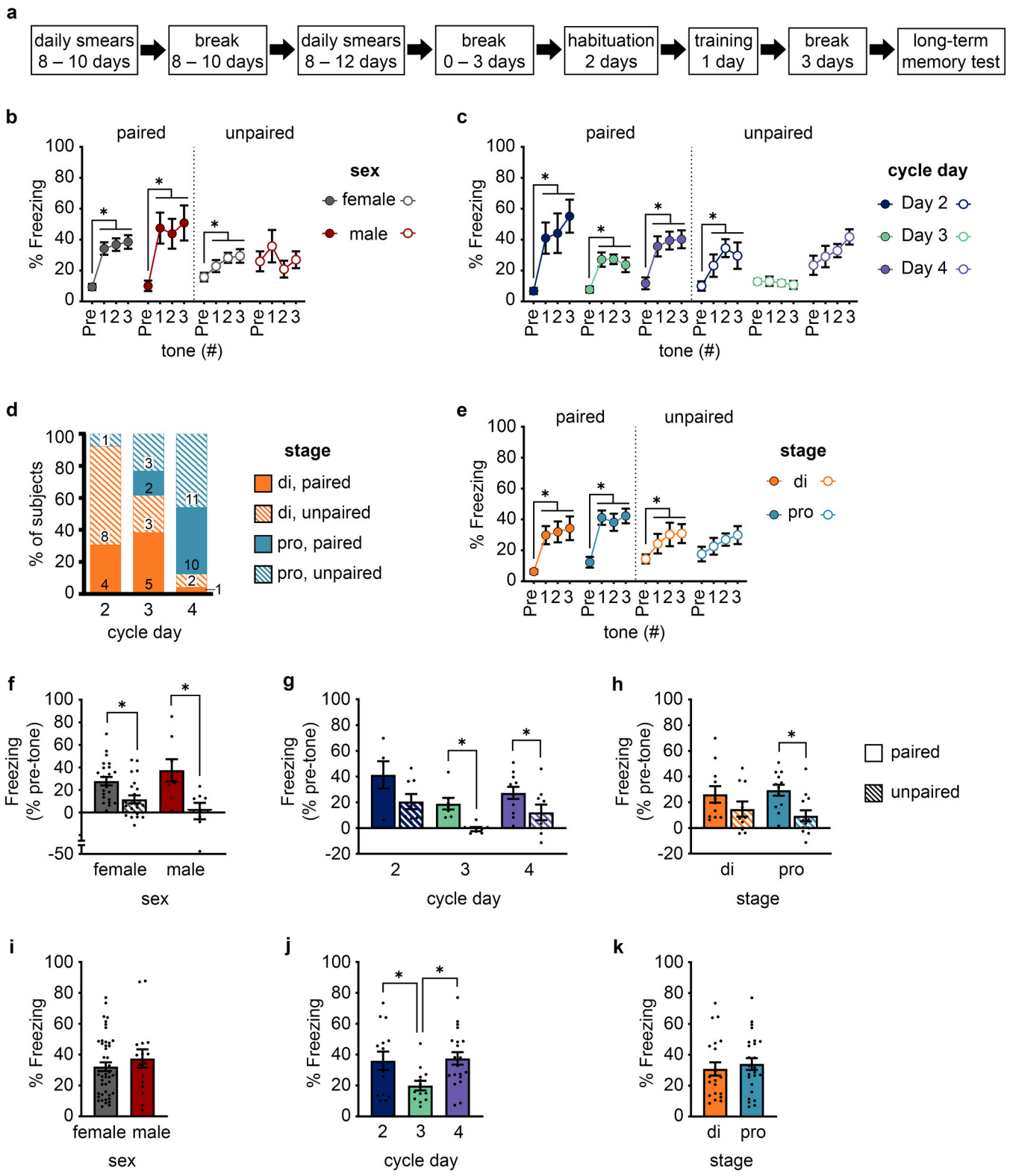


Fig. 6. Track-by-Day detects estrous cycle effects on Pavlovian fear conditioning. (a) Experiment timeline. (b) Freezing before the first tone presentation (Pre) and during the three test tones during the long-term memory test in males and females. Freezing to the tone relative to the pre-tone period was higher in both sexes after paired training, and in females after paired training. (c) Tone freezing was higher than pre-tone after paired training on all cycle days, and after unpaired training on Day 2. (d) Distribution of female subjects by traditional cytology stage and cycle day. (e) Tone freezing was higher in diestrus and proestrus females regardless of training. (f-h) The difference between tone and pre-tone freezing was greater in the paired group versus the unpaired group in both sexes (f), on Day 3 and 4, but not Day 2 (g), and during proestrus but not diestrus (h). (i) Total freezing during the tone for both training groups was not different between males and females. (j-k) Among females, total freezing during the tone in both training groups was lower on Day 3 than on Days 2 and 4 (j), but did not differ between diestrus and proestrus (k). * $p < 0.05$.

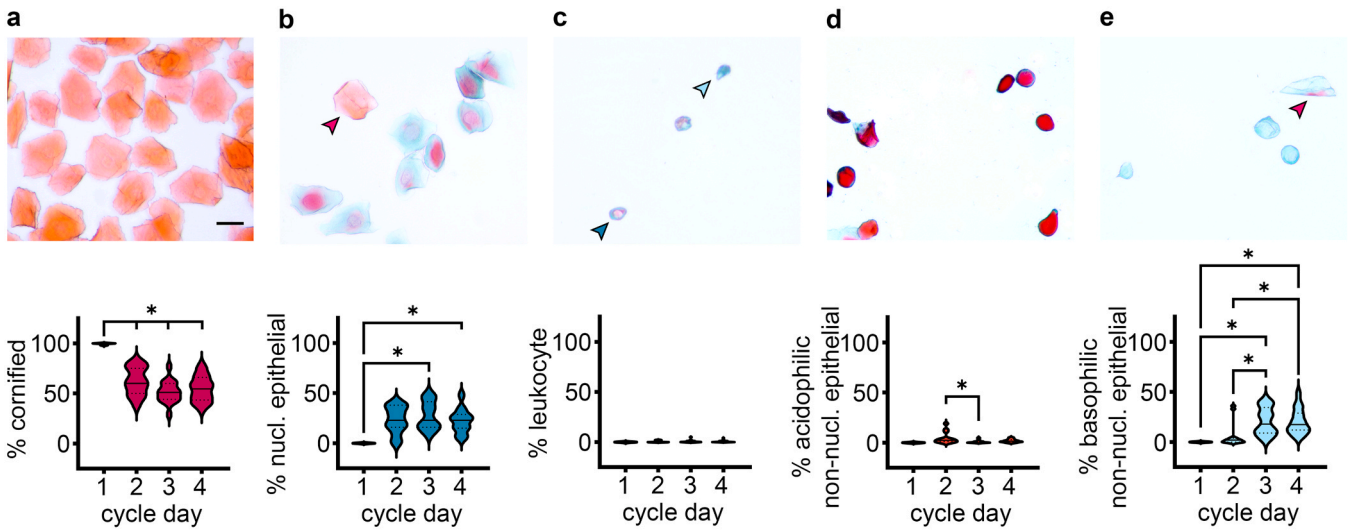


Fig. 7. Quantitative cytology using Shorr stain. (a-e) Appearance of Shorr-stained cells and their percentage on the post-testing smears from the 60 rats in the behavior experiment: (a) cornified; (b) nucleated epithelial cells, with a stray cornified cell (arrow); (c) a leukocyte near a small nucleated epithelial cell (dark arrow) and a basophilic non-nucleated epithelial cell (light arrow); (d) acidophilic non-nucleated epithelial cells; (e) basophilic non-nucleated epithelial cells with a stray basophilic cornified cell (arrow), distinguished by its shape. Scale: 25 μm. * $p < 0.05$.

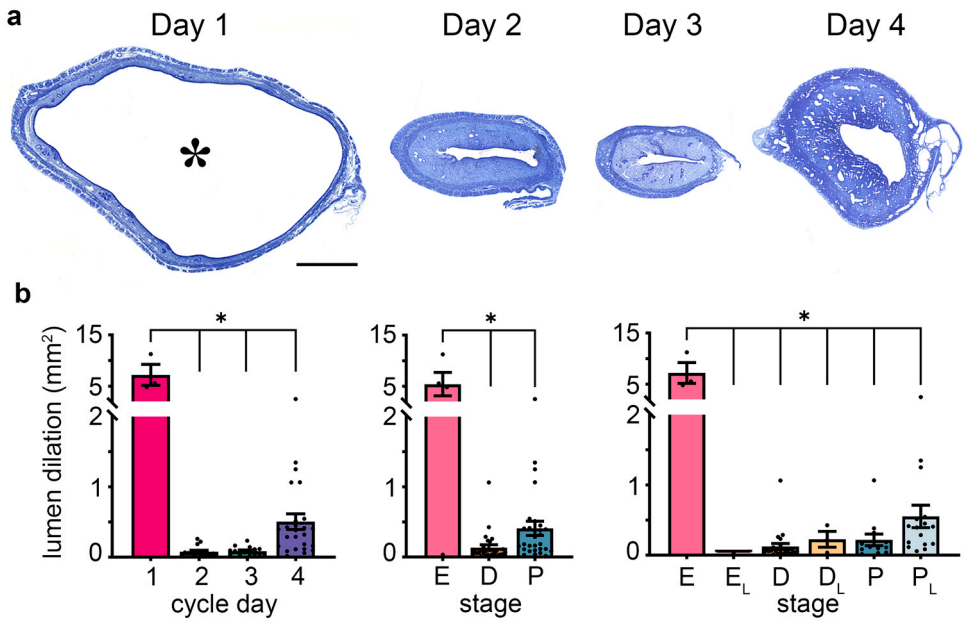


Fig. 8. Gross uterine histology across the cycle. (a) Representative uterine cross sections from each cycle day. The lumen of the Day 1 sample is indicated with a large asterisk. (b) Extreme lumen dilation differentiates Day 1/estrus histology from the rest of the cycle. Scale: 1 μm. * $p < 0.00005$.

associated with each of Days 2, 3, and 4. Mitotic figures in the luminal epithelium (Fig. 9a) were usually present on Days 3 and 4 but never on Day 2. In contrast, luminal mitosis did not differ between diestrus and proestrus or among the transitional stages (Fig. 9b). The presence of glands in the endometrium (Fig. 9c) was also lower on Day 2 versus Days 3 and 4 but was constant between stages (Fig. 9d). Mitotic figures were present in the endometrial glands (Fig. 9e), and their density distinguished Day 3 from the other days but did not differ between stages (Fig. 9f). Finally, the cross-sectional area of the endometrium (Fig. 9g) was greater on Day 4 than Days 2 and 3 (Fig. 9h). This was the only histological measure that differed between stages, being greater in proestrus than diestrus, and greater in late proestrus than early proestrus. Given that two-thirds of proestrus and all late proestrus cases occurred on Day 4 (Supplemental Figure 7–1), this is not surprising.

4. Discussion

The historical practice of using only male subjects has limited our understanding of the female brain, and therefore of the brain in general, and there is a need for more high-quality data from female subjects in most areas of neuroscience. A recent study noted that of research articles that used only males, half cited concerns about variability arising from hormone fluctuations as the reason [60]. Although female subjects are not overall more variable than males [8,57], there is evidence that sex hormones influence a range of brain functions, including learning and memory [28]. It may not be necessary to treat hormonal changes as a potential confound, but it is not unreasonable in areas such as fear- and anxiety-related behaviors, where the literature on estrous cycle effects is inconsistent [22,41]. Serum testosterone conveniently follows a

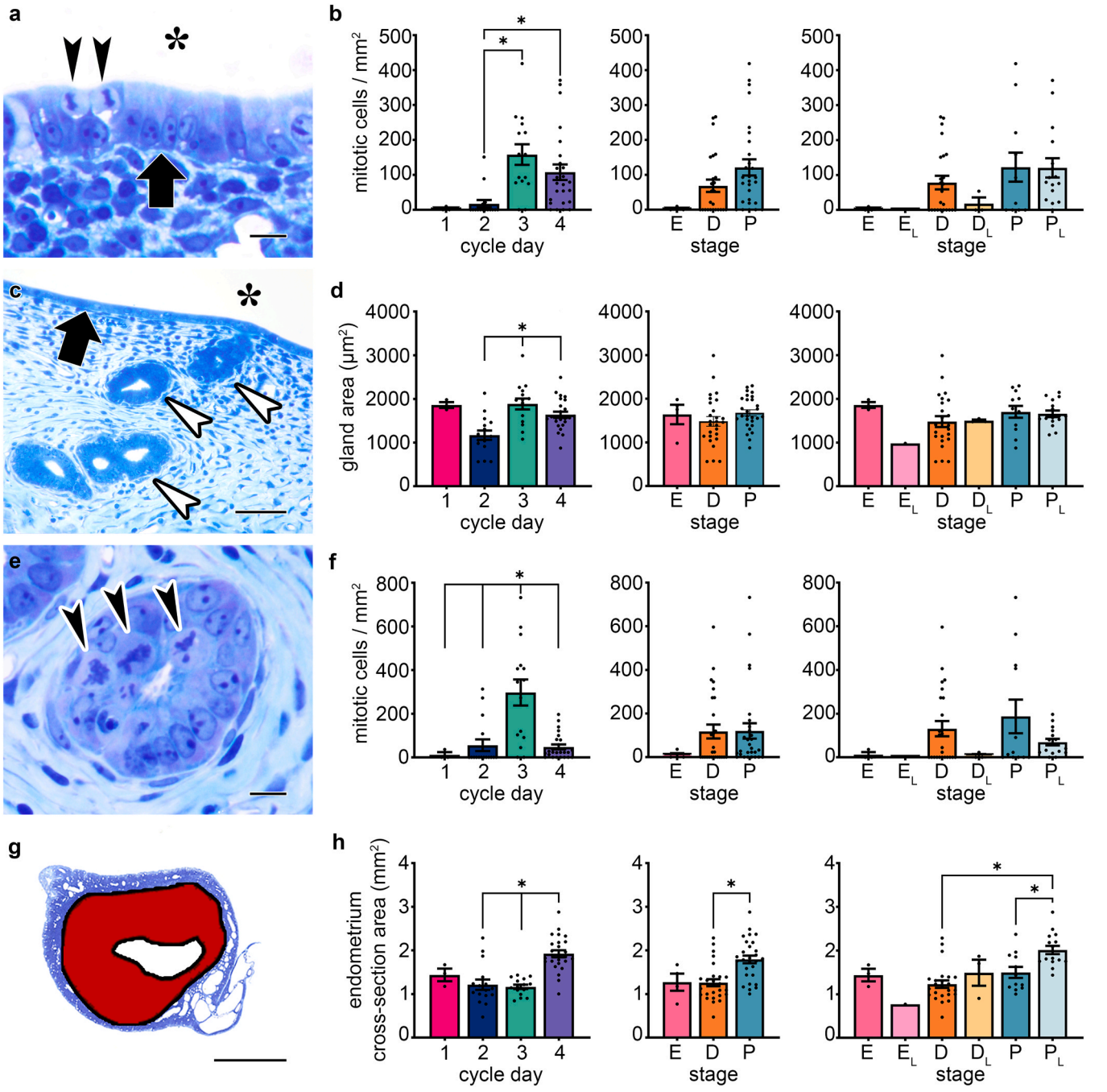


Fig. 9. Cycle days correlate with changes in uterine histology. (a) Mitotic cells (arrowheads) in the luminal epithelium (arrow), which lines the lumen (large asterisk). (b) There were fewer luminal mitotic cells on Day 2 versus Days 3 and 4, but no differences between the stages. (c) Endometrial glands (arrowheads) below the lumen (asterisk) and luminal epithelium (arrow). (d) Endometrial glands covered less area on Day 2 versus Days 3 and 4, but the stages did not differ. (e) Mitotic cells (arrowheads) in an endometrial gland. (f) There were more mitotic cells in endometrial glands on Day 3 versus other days, but no differences between stages. (g) Uterine cross section with endometrium shown by red fill. (h) The cross-sectional area of the endometrium is higher on Day 4 than on Days 2 and 3, and is also higher in proestrus than in diestrus and in late proestrus versus early proestrus and diestrus. Scale bar = 25 µm (a, c, e), 1 µm (g). E: estrus, E_L: late estrus, D: diestrus, D_L: late diestrus, P: proestrus, P_L: late proestrus. * $p < 0.05$.

circadian pattern in male rats [1,48], so controlling for time of day indirectly controls for hormone status. Instead of monitoring the estrous cycle to control for female hormones, however, it is common to avoid female subjects [35,43,74] or ovariectomize and study them under the influence of synthetic hormones [58].

The former may reflect the inaccessibility of tracking methods, and both reflect a lack of confidence in the reliability of the cycle itself.

The first description of changes in vaginal cytology during the rat estrous cycle was published 100 years ago [40]. Cytology was divided into stages according to a four-stage model of mammalian reproduction, which itself was based on observations of reproductive physiology and behavior across many species [33]. It was not until decades later that the hormone cycle was measured in detail and the effects of steroid hormones on the vaginal epithelium were investigated. Several studies have monitored serum levels of ovulation-related hormones across the rat estrous cycle [15,51,70], and two features of the cycle have clear practical implications for cycle tracking. First, hormone levels follow a stereotyped pattern that is entrained to the light-dark cycle, meaning

that cycle progression can be inferred if a single time point can be identified. Second, the major hormonal changes – a surge in 17β-estradiol followed by sharp, nearly simultaneous peaks of progesterone, luteinizing hormone, follicle-stimulating hormone, and prolactin – occur in the 24 h preceding ovulation, and the traditional cytology stages do not correspond to hormone states. Since stage assignments – with the exception of estrus – are effectively arbitrary with respect to hormones, inconsistency between studies that use cytology to stratify groups is unsurprising. Dividing subjects on Day 2 between metestrus and diestrus, for example, creates two groups with essentially identical hormone levels, whereas the cytology does not reflect the rapid hormonal changes on Day 4 in real time and subjects assigned to diestrus or proestrus on Day 4 are unlikely to differ systematically.

Our data show that the most reliably cyclic feature of cytology is the cornified peak and that variability in cytology increases as the cycle progresses, consistent with the relationship between cytology and the hormone peaks becoming more indirect. Our observations are also consistent with cytology guidelines that describe nuanced stage transitions and broad ranges for stage durations. Timing by the cornified peak alone, we found that the vast majority of cycles lasted four days and that almost all rats had regular cycles. This counters the common perception that cycles are often irregular, or, as stated in the title of a recent guide to staging [61], that there is "no such thing as a normal cycle." Since it is common to exclude subjects without regular cycles, differences in exclusion criteria could explain some conflicting results in the literature.

Uterine histology is a more direct reflection of cycle progression than vaginal cytology, as uterine tissue responds more rapidly to 17β-estradiol, and there is evidence that unlike in the vaginal epithelium, the response is dose-dependent [13,31]. Descriptions of changes in uterine histology across the cycle are generally consistent: the lumen dilates on the day of ovulation, accompanied by necrosis of the epithelium and glands[23,64,73]. Water is then expelled from the lumen, and necrosis slows while mitosis resumes, restoring the mature epithelial and gland structures. Our data reflect this temporal progression. The cornified peak in vaginal cytology corresponded 1:1 with profound lumen dilation, and mitosis followed by uterine enlargement differentiated the ensuing days. Only uterine size differed between traditional stages, and variability was higher between stages than between days.

Similarly, comparing across cycle days detected cycle-related differences in cued Pavlovian fear conditioning that were not apparent between cytology stages (Fig. 6), which highlights the utility and sensitivity of the Track-by-Day method. Studies examining the effects of sex and estrous stage on Pavlovian conditioning are inconsistent, with some reporting no differences and others reporting differences in opposing directions [22,41,58]. Consistent with other studies performed in the dark period [11,12,21], we found that subjects of both sexes and all estrous stages and days displayed tone-evoked freezing after paired training. Like another study [12], we looked for darting [32] but found that it was unusual and not systematic; it is possible that this behavior is more common during the light period. Unpaired training produced non-associative freezing in females on Day 2, while Day 3 subjects displayed low freezing regardless of training. When the cycle was not accounted for at all and when traditional staging was used, non-associative freezing appeared to occur in females generally as opposed to in a subset of subjects. Meanwhile, the freezing effect on Day 3 was completely obscured when subjects were grouped by stage. Neither non-associative freezing nor lowered freezing can be attributed to differences in shock sensitivity, which did not change across the cycle. Non-associative freezing is likely due to pseudoconditioning, which is mechanistically distinct from associative learning and thought to be a non-specific arousal effect[2,24,36]. Interestingly, a recent study of male and female rats in a naturalistic foraging setting found that a tone stimulus evoked non-associative freezing but not Pavlovian conditioning [79]. Pseudoconditioning in Day 2 females could thus represent ethologically-relevant vigilance or defensive behavior that is reduced later in the cycle. Lowered freezing on Day 3 does not indicate impaired learning, as associative freezing to the tone was preserved, but instead a difference in behavioral expression.

Grouping by cycle day in the dark period yields groups with presumed low hormones (Day 2), high progesterone (Day 3), or high 17β-estradiol (Day 4), so unlike staging, Track-by-Day allows interpretation in the context of hormone levels. Endogenous and exogenous progesterone and 17β-estradiol have been reported to be anxiolytic in females [58]. In the elevated plus maze, progesterone is associated with more open arm entries as well as more open arm time in the elevated plus maze[10,30,50], whereas acutely high 17β-estradiol has been associated with more open arm time but not more entries [45,72]. Although freezing in a Pavlovian conditioning paradigm is not considered a measure of anxiety, these effects are interesting in the context of our data. If both steroids reduce anxiety responses, the non-associative freezing on Day 2 could reflect higher vigilance in the low-hormone state. If progesterone drives more active behavior in response to threats than 17β-estradiol, that could explain the overall low freezing on Day 3.

Tracking data from 328 cycles in 94 rats demonstrated regular cycling across conditions and subjects. Synchrony between cage-mate pairs occurred at chance levels, as was also reported in a quantitative study of cycle timing in group-housed rats [65]. The majority of cycles (85%) lasted four days, and only four rats failed to have any four-day cycles. Of these, one had no discernable cycles and was confirmed to be pseudopregnant by uterine histology, while the other three had five-day cycles. Most five-day cycles occurred sporadically in subjects that also had multiple four-day cycles, and there was only one apparent instance of two consecutive five-day cycles. These results differ from much of the older literature, which reports that four-day and five-day cycles are characteristic of individual rats. It is possible that the Sprague-Dawley rats we used are especially prone to four-day cycles; some of the classic work on cycle length regulation made use of two sub-strains of Osborne-Mendel rats bred to consistently exhibit either four-day or five-day cycles [25,26], indicating a genetic contribution to cycle length.

Quantitative data on four- and five-day cyclers is scarce, but a large study of Sprague-Dawley rats reported that 247 out of more than 500 subjects had two consecutive cycles of the same length, and of those 247 roughly equal numbers had four- and five-day cycles [66]. The failure of more than half the subjects to have two same-length cycles in a row could be due to strict stage-based cycle definitions, but the overall rate of about 25% of subjects having two consecutive five-day cycles is much higher than what we observed. It is possible that their rats, like ours, exhibited mixed cycle lengths but with a higher proportion of five-day cycles – if every cycle in every rat has an equal likelihood of lasting four or five days, pairs of consecutive cycles would yield the reported ratio of 25% four-day, 25% five-day, and 50% "irregular" cyclers. A later study by the same author, however, noted that a third consecutive cycle of the same length occurred in 75% of four-day and 85% of five-day cyclers, indicating a degree of individual bias in cycle length [51]. That study quantified serum hormones in terminal blood samples from four- and five-day cyclers (defined by the two previous consecutive cycles) and found differences mainly on Days 4 and 1/5 (data on progesterone and 17β-estradiol are adapted in Fig. 1a). Rats whose uterine and oviduct histology did not confirm the expected point of the ovulatory cycle were excluded, meaning that the data represent only cycles of the specified length as opposed to consistent patterns in individuals prone to a given cycle length (i.e. rats with prior five-day cycles do not necessarily have delayed hormone peaks on four-day cycles).

As illustrated in Fig. 1a, the serum hormone patterns in four-day cycles reported by Nequin and colleagues (1979) correspond to those reported by two other studies [15,70]. Both of those used only rats with three consecutive four-day cycles, and although the numbers are not reported, presumably rats with consistent five-day cycles were not frequent enough for analysis. All three studies used Sprague-Dawley rats, and the reason for the higher incidence of five-day cycles in the

studies from Schwartz and colleagues [51,66] is unclear. Subjects in those studies were housed on a 14:10 h light:dark cycle, a condition under which we observed only four-day cycles in our small cohort of 12 rats. We collected tracking data from subjects housed on 12:12 or 14:10 h light:dark cycles, subjects housed in active breeding colonies or with only female rats, and from subjects before and during the stress of daily blood draws. Across all conditions, four-day cycles were the rule and five-day cycles occurred seemingly at random. Although we cannot rule out unknown environmental factors, the most likely explanation for our failure to see a substantial number of five-day cycles is a congenital tendency to four-day cycles. All of the subjects in this study were obtained from a single vendor, but in subsequent experiments we have had similar results with Sprague-Dawley rats from a different vendor. Given the potential for genetic variation in cycle length, it is advisable to observe patterns over a few cycles when beginning work with a new strain or vendor.

Cycle length and regularity are also reported to vary in rats younger than one month and older than ten months [42,46,77], so our results should not be assumed to apply before or after adulthood, when reproductive function is cyclic and stable. Sprague-Dawley rats are also an especially docile strain and do not show behavioral signs of stress during swabbing, but is possible that strains which require acclimation to handling, such as Long-Evans, will experience stress during initial attempts at smear collection. Cycle length was not affected by the stress of daily blood draws in our study, consistent with a report of four-day cycle length persisting during daily restraint stress in Sprague-Dawley rats [12]. Other strains may be more sensitive, however, and to avoid other effects of stress it would be prudent to ensure that subjects are habituated to swabbing before beginning experiments.

In contrast to traditional staging, the Track-by-Day method does not require specialized training in cytology interpretation, making it accessible to a wide community of researchers. It also inherently facilitates standardization across laboratories; regardless of whether traditional stages are given, studies that report the cycle day and time of data collection can be easily compared. In classic endocrinology literature [15,51,70] the names of the four stages are used to refer to cycle days as opposed to cytology, and this convention is followed by at least one neuroscience group [11,12]. Because of the longstanding association between the stage names and cytology states of variable duration, we propose that the cycle be reported in days, as it is for golden hamsters [38,47,52,56]. This will avoid ambiguity and allow traditional stages to be reported in parallel if desired. Ideally, by making the estrous cycle accessible the Track-by-Day method will encourage more researchers to work with female subjects and improve our understanding of how the estrous cycle does, or does not, influence behavior.

5. Recommendations for the track-by-day method

Planning: Track smears for twelve consecutive days before beginning an experiment. This will acclimate the subjects (and experimenters) to the procedure and provide two full cycles of data.

Sample collection: Collect smears once per day, ideally between the mid-dark period and mid-light period. Contrast smears with H&E or Shorr stain.

Interpretation: Identify peak smears containing only cornified cells, especially in clumps and sheets.

Tracking: Define fully-cornified smears as Day 1, and assign non-peak smears to the corresponding cycle days. If two consecutive peak cornified smears are observed, the final smear before a non-peak smear is Day 1. This ensures that the cycle is timed from the day of ovulation.

Reporting: Subjects can be grouped for analysis by cycle days instead of or in addition to traditional stages, but whether stages are reported or not, the cycle day relative to the cornified peak and the exact time that experiments were performed should be reported to allow standardized comparisons between studies.

Funding

This work was supported by NSF 2014862, the Connecticut Institute for the Brain and Cognitive Sciences, and the Arnold and Mabel Beckman Foundation.

Author Statement

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

Declaration of Competing Interest

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

Data Availability

Data will be made available on request.

Acknowledgements

We are grateful to Maritza Abril, Zachary Deane, Jairo Orea, and Jerry Vasquez Jr. for help with some of the experiments, Teng Long and Joy Li for help with behavioral analysis, and Chris Cain, Michael Goard, Alex Jackson, John Redden, Weiling Yin, and Rebecca Tripp for helpful comments on the manuscript. This work was performed in part at the Bioscience Electron Microscopy Laboratory of the University of Connecticut.

Author Contributions

Conceived and designed the study: LO and GR; performed experiments: GR, AE, MK, and MB; analyzed data: GR, MK, LO; wrote the paper: GR and LO.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.bbr.2024.114860.

References

[1] A.Z. Baburski, S.J. Sokanovic, M.M. Bjelic, S.M. Radovic, S.A. Andric, T.S. Kostic, Circadian rhythm of the Leydig cells endocrine function is attenuated during aging, Exp. Gerontol. 73 (2016) 5–13.
[2] J.S. Bakin, B. Lepan, N.M. Weinberger, Sensitization induced receptive field plasticity in the auditory cortex is independent of CS-modality, Brain Res. 577 (1992) 226–235.
[3] D.A. Bangasser, A. Cuarenta, Sex differences in anxiety and depression: circuits and mechanisms, Nat. Rev. Neurosci. 22 (2021) 674–684.
[4] T.E. Barker, B.E. Walker, Initiation of irreversible differentiation in vaginal epithelium, Anat. Rec. 154 (1966) 149–159.
[5] L. Bartos, Vaginal impedance measurement used for mating in the rat, Lab Anim. 11 (1977) 53–55.
[6] W.W. Beatty, P.A. Beatty, Hormonal determinants of sex differences in avoidance behavior and reactivity to electric shock in the rat, J. Comp. Physiol. Psychol. 73 (1970) 446–455.
[7] J.B. Becker, A.P. Arnold, K.J. Berkley, J.D. Blaustein, L.A. Eckel, E. Hampson, J. P. Herman, S. Marts, W. Sadee, M. Steiner, et al., Strategies and methods for research on sex differences in brain and behavior, Endocrinology 146 (2005) 1650–1673.
[8] J.B. Becker, B.J. Prendergast, J.W. Liang, Female rats are not more variable than male rats: a meta-analysis of neuroscience studies, Biol. Sex. Differ. 7 (2016).
[9] A.K. Beery, Inclusion of females does not increase variability in rodent research studies, Curr. Opin. Behav. Sci. 23 (2018) 143–149.
[10] D. Bitran, R.H. Purdy, C.K. Kellog, Anxiolytic effect of progesterone is associated with increases in cortical allopregnanolone and GABAA receptor function, Pharm. Biochem. Behav. 45 (1993) 423–428.

[11] S.R. Blume, M. Freedberg, J.E. Vantrease, R. Chan, M. Padival, M.J. Record, M. R. De Joseph, J.H. Urban, J.A. Rosenkranz, Sex- And estrus-dependent differences in rat basolateral amygdala, J. Neurosci. 37 (2017) 10567–10586.

[12] S.R. Blume, M. Padival, J.H. Urban, J.A. Rosenkranz, Disruptive effects of repeated stress on basolateral amygdala neurons and fear behavior across the estrous cycle in rats, Sci. Rep. 9 (1) (2019) 18.

[13] R.M. Brenner, N.B. West, Hormonal regulation of the reproductive tract in female mammals, Annu. Rev. Physiol. 37 (1975) 273–302.

[14] D.L. Buchanan, T. Kurita, J.A. Taylor, D.B. Lubahn, G.R. Cunha, P.S. Cooke, Role of stromal and epithelial estrogen receptors in vaginal epithelial proliferation, stratification, and cornification, Endocrinology 139 (1998) 4345–4352.

[15] R.L. Butcher, W.E. Collins, N.W. Fugo, Plasma concentration of LH, FSH, prolactin, progesterone and estradiol-17beta throughout the 4-day estrous cycle of the rat, Endocrinology 94 (1974) 1704–1708.

[16] S.L. Byers, M.V. Wiles, S.L. Dunn, R.A. Taft, Mouse estrous cycle identification tool and images, PLoS One 7 (2012) 35538.

[17] K.L. Chesney, C. Chang, E.C. Bryda, Using vaginal impedance measurement to identify proestrus in rats given luteinizing hormone releasing hormone (lhrh) agonist, J. Am. Assoc. Lab Anim. Sci. 59 (2020) 282–287.

[18] E. Choleris, L.A.M. Galea, F. Sohrabji, K.M. Frick, Sex differences in the brain: implications for behavioral and biomedical research, Neurosci. Biobehav Rev. 85 (2018) 126–145.

[19] P.S. Cooke, T.E. Spencer, F.F. Bartol, K. Hayashi, Uterine glands: development, function and experimental model systems, Mol. Hum. Reprod. 19 (2013) 547–558.

[20] M.C. Cora, L. Kooistra, G. Travlos, Vaginal cytology of the laboratory rat and mouse: review and criteria for the staging of the estrous cycle using stained vaginal smears, Toxicol. Pathol. 43 (2015) 776–793.

[21] R. Cossio, M.B. Carreira, C.E. Vásquez, G.B. Britton, Sex differences and estrous cycle effects on foreground contextual fear conditioning, Physiol. Behav. 163 (2016) 305–311.

[22] H.L.L. Day, C.W. Stevenson, The neurobiological basis of sex differences in learned fear and its inhibition, Eur. J. Neurosci. 52 (2020) 2466–2486.

[23] D. Dixon, R. Alison, U. Bach, K. Colman, G.L. Foley, J.H. Harleman, R. Haworth, R. Herbert, A. Heuser, G. Long, et al., Nonproliferative and proliferative lesions of the rat and mouse female reproductive system, J. Toxicol. Pathol. 27 (2014).

[24] M.T. Erickson, E.T. Walter, Differential expression of pseudoconditioning and sensitization by siphon responses in Aplysia: Novel response selection after training, J. Neurosci. 8 (1988) 3000–3010.

[25] J.W. Everett, Progesterone and estrogen in the experimental control of ovulation time and other features of the estrous cycle in the rat, Endocrinology 43 (1948) 389–405.

[26] J.W. Everett, Pseudopregnancy in the rat from brief treatment with progesterone: effect of isolation, Nature 198 (1963) 695–696.

[27] J.C. Fiala, Reconstruct: a free editor for serial section microscopy, J. Microsc. 218 (2005) 52–61.

[28] A.W. Fleischer, K.M. Frick, New perspectives on sex differences in learning and memory, Trends Endocrinol. Metab. 34 (2023).

[29] K.M. Frick, J. Kim, Mechanisms underlying the rapid effects of estradiol and progesterone on hippocampal memory consolidation in female rodents, Horm. Behav. 104 (2018) 100–110.

[30] C.A. Frye, S.M. Petralia, M.E. Rhodes, Estrous cycle and sex differences in performance on anxiety tasks coincide with increases in hippocampal progesterone and 3α,5α-THP, Pharm. Biochem. Behav. 67 (2000) 587–596.

[31] P. Galand, F. Leroy, J. Chrétien, Effect of oestradiol on cell proliferation and histological changes in the uterus and vagina of mice, J. Endocrinol. 49 (1971) 243–252.

[32] T.M. Gruene, K. Flick, A. Stefano, S.D. Shea, R.M. Shansky, Sexually divergent expression of active and passive conditioned fear responses in rats, Elife 4 (2015) 1–9.

[33] W. Heape, The "sexual season" of mammals and the relation of the "pro-estrum" to menstruation, Q J. Microsc. Sci. 44 (1900) 1–70.

[34] C.H. Hubscher, D.L. Brooks, J.R. Johnson, A quantitative method for assessing stages of the rat estrous cycle, Biotech. Histochem 80 (2005) 79–87.

[35] R.N. Hughes, Sex still matters: has the prevalence of male-only studies of drug effects on rodent behaviour changed during the past decade? Behav. Pharm. 30 (2019) 95–99.

[36] J. Iwata, K. Chida, J.E. LeDoux, Cardiovascular responses elicited by stimulation of neurons in the central amygdaloid nucleus in awake but not anesthetized rats resemble conditioned emotional responses, Brain Res. 418 (1987) 183–188.

[37] L.M. Jaramillo, I.B. Balcazar, C. Duran, Using vaginal wall impedance to determine estrous cycle phase in Lewis rats, Lab. Anim. 41 (2012) 122–128.

[38] H. Kon, H. Kishi, K.Y. Arai, M. Shinoda, G. Watanabe, K. Taya, The effects of prolactin and gonadotropin on luteal function and morphology in the cyclic golden hamster, J. Reprod. Dev. 54 (2008) 418–423.

[39] F. Leroy, P. Galand, J. Chrétien, The mitogenic action of ovarian hormones on the uterine and the vaginal epithelium during the oestrous cycle in the rat: a radioautographic study, J. Endocrinol. 45 (1969) 441–447.

[40] J.A. Long, H.M. Evans, The Oestrous Cycle in the Rat and its Associated Phenomena, University of California Press, Berkeley, Calif, 1922.

[41] T.A. Lovick, H. Zangrossi, Effect of estrous cycle on behavior of females in rodent tests of anxiety, Front. Psychiatry 12 (2021).

[42] K.H. Lu, B.R. Hopper, T.M. Vargo, S.S.C. Yen, Chronological changes in sex steroid, gonadotropin and prolactin secretion in aging female rats displaying different reproductive states1, Biol. Reprod. 21 (1979) 193–203.

[43] G.M. Mamlouk, D.M. Dorris, L.R. Barrett, J. Meitzen, Sex bias and omission in neuroscience research is influenced by research model and journal, but not reported NIH funding, Front. Neuroendocr. (2020) **57**.

[44] F.K. Marcondes, F.J. Bianchi, A.P. Tanno, Determination of the estrous cycle phases of rats: some helpful considerations, Braz. J. Biol. 62 (2002) 609–614.

[45] F.K. Marcondes, K.J. Miguel, L.L. Melo, R.C. Spadari-Bratfisch, Estrous cycle influences the response of female rats in the elevated plus-maze test, Physiol. Behav. 74 (2001) 435–440.

[46] M. McClintock, J. LeFevre, Reproductive senescence in female rats: a longitudinal study of individual differences in estrous cycles and behavior, Behav. *Biol. Reprod.* 38 (1988) 780–789.

[47] J.T. McCormack, M.G. Friederichs, S.D. Limback, G.S. Greenwald, Apoptosis during spontaneous luteolysis in the cyclic golden hamster, Biochem. Morphol. Evid. *Biol. Reprod.* 58 (1998) 255–260.

[48] E.J. Mock, H.W. Norton, A.I. Frankel, Daily rhythmicity of serum testosterone concentration in the male laboratory rat, Endocrinology 103 (1978) 1111–1121.

[49] G.S. Montes, E.H. Luque, Effects of ovarian steroids on vaginal smears in the rat, Cells Tissues Organs 133 (1988) 192–199.

[50] S. Mora, N. Dussaubat, G. Díaz-Véliz, Effects of the estrous cycle and ovarian hormones on behavioral indices of anxiety in female rats, Psychoneuroendocrinology 21 (1996) 609–620.

[51] L.G. Nequin, J. Alvarez, N.B. Schwartz, Measurement of serum steroid and gonadotropin levels and uterine and ovarian variables throughout 4 day and 5 day estrous cycles in the Rat1, Biol. Reprod. 20 (1979) 659–670.

[52] M. Orsini, The external vaginal phenomena characterizing the stages of the estrous cycle, pregnancy, pseudopregnancy, lactation, and the anestrous hamster, Mesocricetus auratus Waterhouse, Proc. Anim. Care Panel 11 (1961) 193–206.

[53] L.E. Ostroff, C.K. Cain, J. Bedont, M.H. Monfils, J.E. LeDoux, Fear and safety learning differentially affect synapse size and dendritic translation in the lateral amygdala, Proc. Natl. Acad. Sci. USA 107 (2010) 9418–9423.

[54] Paccola C.C., Resende C.G., Stumpp T., Miraglia S.M., Cipriano I. 2018. *The rat estrous cycle revisited: a quantitative and qualitative analysis*. Colégio Brasileiro de Reprodução Animal.

[55] R.V. Pearse, T.L. Young-Pearse, Lost in translational biology: Understanding sex differences to inform studies of diseases of the nervous system, Brain Res. 1722 (2019) 146352.

[56] B.J. Prendergast, A.K. Beery, M.J. Paul, I. Zucker, Enhancement and suppression of ultradian and circadian rhythms across the female hamster reproductive cycle, J. Biol. Rhythms 27 (2012) 246–256.

[57] B.J. Prendergast, K.G. Onishi, I. Zucker, Female mice liberated for inclusion in neuroscience and biomedical research, Neurosci. Biobehav. Rev. 40 (2014) 1–5.

[58] M.E. Price, B.A. McCool, Structural, functional, and behavioral significance of sex and gonadal hormones in the basolateral amygdala: A review of preclinical literature, Alcohol 98 (2022) 25–41.

[59] S.D. Ramos, J.M. Lee, J.D. Peuler, An inexpensive meter to measure differences in electrical resistance in the rat vagina during the ovarian cycle, J. Appl. Physiol. 91 (2001) 667–670.

[60] R.K. Rechlin, T.F.L. Splinter, T.E. Hodges, A.Y. Albert, L.A.M. Galea, An analysis of neuroscience and psychiatry papers published from 2009 and 2019 outlines opportunities for increasing discovery of sex differences, Nat. Commun. 13 (2022).

[61] H. Robert, L. Ferguson, O. Reins, T. Greco, M.L. Prins, M. Folkerts, Rodent estrous cycle monitoring utilizing vaginal lavage: no such thing as a normal cycle, *J. Vis. Exp.* (2021).

[62] D.R. Rubinow, P.J. Schmidt, Sex differences and the neurobiology of affective disorders, Neuropsychopharmacology 44 (2019) 111–128.

[63] K. Sano, S. Matsuda, S. Tohyama, D. Komura, E. Shimizu, C. Sutoh, Deep learning-based classification of the mouse estrous cycle stages, Sci. Rep. 10 (2020).

[64] J. Sato, M. Nasu, M. Tsuchitani, Comparative histopathology of the estrous or menstrual cycle in laboratory animals, J. Toxicol. Pathol. 29 (2016) 155–162.

[65] J.C. Schank, Do Norway rats (Rattus norvegicus) synchronize their estrous cycles? Physiol. Behav. 72 (2001) 129–139.

[66] N.B. Schwartz, Acute effects of ovariectomy on pituitary LH, uterine weight, and vaginal cornification, Am. J. Physiol. 207 (1964) 1251–1259.

[67] R.M. Shansky, Are hormones a "female problem" for animal research? Outdated gender stereotypes are influencing experimental design in laboratory animals, Science 364 (2019) 825–826.

[68] Shorr E. 1941. A New Technic for Staining Vaginal Smears: III, a Single Differential Stain Author (s): Ephraim Shorr Published by: American Association for the Advancement of Science Stable URL: ⟨https://www.jstor.org/stable/1667807⟩. *Science* **94**: 545–546.

[69] S.J. Singletary, A.J. Kirsch, J. Watson, B.O. Karim, D.L. Huso, P.D. Hurn, S. J. Murphy, Lack of correlation of vaginal impedance measurements with hormone levels in the rat, Conte Top. Lab Anim. Sci. 44 (2005) 37–42.

[70] M.S. Smith, M.E. Freeman, J.D. Neill, The control of progesterone secretion during the estrous cycle and early pseudopregnancy in the rat: prolactin, gonadotropin and steroid levels associated with rescue of the corpus luteum of pseudopregnancy, Endocrinology 96 (1975) 219–226.

[71] C.R. Stockard, G.N. Papanicolaou, The existence of a typical oestrous cycle in the guinea pig with a study of its histological and physiological changes, Am. J. Anat. 22 (1917) 225–283.

[72] A.A. Walf, C.A. Frye, Antianxiety and antidepressive behavior produced by physiological estradiol regimen may be modulated by hypothalamic-pituitary-adrenal axis activity, Neuropsychopharmacology 30 (2005) 1288–1301.

[73] F.R. Westwood, The female rat reproductive cycle: a practical histological guide to staging, Toxicol. Pathol. 36 (2008) 375–384.

[74] Woitowich N.C., Beery A., Woodruff T. 2020. A 10-year follow-up study of sex inclusion in the biological sciences. 1–8.
[75] N.S. Wolcott, K.K. Sit, G. Raimondi, T. Hodges, R.M. Shansky, L.A.M. Galea, L. E. Ostroff, M.J. Goard, Automated classification of estrous stage in rodents using deep learning, bioRxiv (2022), 2022.03.09.483678.
[76] T. Yener, A. Turkkani Tunc, H. Aslan, H. Aytan, A.C. Caliskan, Determination of oestrous cycle of the rats by direct examination: how reliable? Anat. Histol. Embryol. 36 (2007) 75–77.
[77] Yin W., Maguire S.M., Pham B., Garcia A.N., Dang N., Liang J., Wolfe A., Hofmann H.A., Gore A.C. 2015. Testing the Critical Window Hypothesis of Timing and Duration of Estradiol Treatment on Hypothalamic Gene Networks in Reproductively Mature and Aging Female Rats. 1–16.
[78] S.R.K. Zajitschek, F. Zajitschek, R. Bonduriansky, R.C. Brooks, W. Cornwell, D. S. Falster, M. Lagisz, J. Mason, A.M. Senior, D.W.A. Noble, et al., Sexual dimorphism in trait variability and its eco-evolutionary and statistical implications, Elife 9 (2020) 1–17.
[79] P.R. Zambetti, B.P. Schuessler, B.E. Lecamp, A. Shin, E.J. Kim, J.J. Kim, Ecological analysis of Pavlovian fear conditioning in rats, Commun. Biol. 5 (1) (2022) 11.