nature portfolio

Peer Review File

Adaptation of hypothalamic CRH neuron responses to stress and pup cues during motherhood

Corresponding Author: Dr Karl Iremonger

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
This interesting study used fibre photometry to investigate PVN CRH cell dynamics in virgin and postpartum mice in response to acute handling stress, a novel object, and pups. Because the literature on this topic concludes that the CRH system is suppressed by lactation, and there are no electrophysiology studies of how this cell group responds to stressors or other stimuli in lactating animals, the surprising result here was almost no differences between virgin and lactating mice (either those with pups or that had pups removed days before) in PVN CRH cell activity to the probably major acute stress of handling or the minor acute stress of a novel object. This was despite these animals showing a peripheral HPA axis response to the stress. Differences in PVN CRH cell firing were found when mice interacted with a pup directly or couldn't get to the pup inside a metal ball. Under those circumstances lactating mice allowed to remain with their pups showed decreased PVN CRH cell activity if they could interact with the test pup, but there was more PVN CRH cell activity if the test pup was inaccessible, which indicates frustration to an unachievable goal. The results are novel and would broaden understanding of how lactation affects PVN CRH cells and probably CRH cells in other brain sites.

1) The longer I thought about these results the less I was so surprised by them. The authors state starting in the abstract that the lack of suppression of CRH cell activity is contrary to expectation. But just the well known decrease in CRH synthesis by those PVN cells plus the effects of lactation on the pituitary wouldn't further require suppression of CRH cell activity rate to reduce the HPA axis response during lactation. So not entirely a surprise? The increase in activity seen in response to the pup in the metal ball is more notable, but because these cells synthesize and release other neurotransmitters, and these experiments didn't examine what they're releasing in any terminal site, I think there's potential for overinterpretating these results. I also think this result isn't completely surprising in the context of the Walker and Woodside work 20 years ago cited by the authors in the discussion showing that even the peripheral HPA axis response can be strongly activated if the stress is in the presence of the pups.

2) There's no consideration of the heterogeneity of PVN CRH cells. Supp Fig 1 seems to show the entire dorsal PVN infected with virus at that level. There's no depiction or description of where the optic implants were located but maybe the recording area is so large and diffuse it doesn't matter. What percentage of the PVN CRH cells in these female mice were infected? For the cells that show fast activity changes to pups, do the data across subjects provide any insight into where they were in the PVN? How are they different from other CRH cells nearby? Does that population project somewhere that's different from other CRH cells that don't respond to pups? Is there any literature on heterogeneity of PVN CRH cells that could help the data interpretation? I think the manuscript would be strengthened if some of these or other questions were answered. Expanding the discussion despite the journal's word limit would also help.

3) I'm wondering whether or not 24 hours separated from pups was long enough to restore the cell activity function of the PVN CRH cells as detectable by the photometry. They cite studies showing that 24 hours is enough to restore the CRH mRNA synthesis in those cells, but could a longer separation have been better for this study? Related to separation in the novel object experiment, the lactating mice have pups taken away I think for a range of 2-4 days before recording. Did this range create heterogeneity in the results? Is there more to dig into the data there? I also wonder why mice remaining with their pups were removed from them 1 hour before recording and whether or not this affected the results.

4) The latencies to contact pups is very long in Supp Fig 1, at about three minutes, and makes me concerned about the behavior of these animals and whether they're somehow stressed at baseline and not responding properly. Was the lack of

group difference in that and the other behaviors expected when it was the very first time the virgins were shown the pups?

5) Could baseline firing of PVN CRH cells differ between virgins and lactating mice and affect the photometry data when the data for each group is relative to different baselines? Maybe this is accounted for in the data analysis in ways I'm not clear about.

6) The authors may want to emphasize in the discussion that it's only the PVN CRH neurons they targeted so the results may not extend to how other groups of CRH cells function. Is
anything known about other CRH cells in other sites during lactation?

7) The authors can specify what CRH cell population they're referring to in each instance. I didn't know when they were talking about their or other studies whether or not they always meant it was the PVN population.

8) Is the X axis on Supp Figure 2 the total number of pup calls? Can trend lines be added to those panels to help make the group differences in the correlations more obvious? I don't know what to make of the data in panel G because that one dataset doesn't have the full distribution of calls compared to the other two. Is that a statistical concern when comparing correlations between groups when the underlying data distributions differ? Maybe convert data to Z-scores? The Y axis on that panel also differs from the others for some reason.

9) Animals paragraph in methods - after pregnancy was confirmed, males were removed but were the females housed singly or group housed? The minimum number of pups was 4 but what was the maximum and do the data from the study show whether the number of pups being raised affects the maternal CRH cells or their behavior? If the range of litter size is large, maybe that needs to be a covariate.

Reviewer #2

(Remarks to the Author)
The manuscript by Kim et al., examines dynamic PVNCRH activity patterns in lactating dams in response to different environmental stressors and during pup retrieval. Both Crh mRNA and cort data are consistent with the literature that PVNCRH/HPA activity is suppressed in postpartum dams. Yet, the investigators have found conditions in which the PVNCRH dynamics are unchanged or enhanced compared to virgin females.

This is an interesting study that would be of interest to a broad readership. However, there are several issues in the manuscript that must be resolved in order for the study to be suitable for publication. These issues/questions are enumerated in detail below. These concerns need to be fully addressed, and the manuscript revised prior to resubmission.

1. The description of the statistics used are insufficient. There are several cases in which use of two-way ANOVAs (in some cases repeated measures, in some cases not) are required in order for the investigators to make their conclusions. Yet it is unclear that they are using the appropriate statistics: Figure 2c,3c,4c, 5d,f,g,h, most of the comparisons in Figure 6. The only place this is indicated is in the Figure legend of Figure 7. For within subject comparisons, the investigators can normalize to baseline and then run one-way ANOVAs. It may also help the reader to show the normalized responses in the main text and raw data in an extended figure for Figure 6 since most of the comparisons appear to be within subject baseline vs. post-stimulus. In some places this is done and in others it is not. It would be good to be consistent in how the data are displayed across figures.
2. Provide technical details of fiber photometry recordings in the methods rather than just referencing the two citations. Specifically, was use of an isosbestic signal (e.g. 405) employed. If not, how were issues like motion artifact corrected for?
3. Normalization of event-related transients is standard; however, normalization of spontaneous transients is problematic. It is better to report the raw frequency and peak amplitude. See Wallace et al., 2025, ACS Chemical Neuroscience.
4. In addition to reporting the cumulative signal (and also AUC would be useful) for the entire 5 or 10 minute duration of the stimulus in Figure 2 and 3, the investigators should isolate and expand the time window of the onset, sustained and offset components. It may be that that by not isolating and analyzing different components of the stimulus presentation, they are missing valuable information regarding how PVNCRH neurons are differentially encoding environmental stressors in virgin vs. lactating females.
5. The methods for collecting corticosterone are absent. It is unclear if the "basal" and "whitenoise" blood is collected within group or in separate sets of mice. If it is the latter, there should not be connecting lines. Seems like it is within subject…in which case, is it tail blood that is being collected?
6. Clarify that what is measured is "area/cell" in the y-axis labeling in 4b.
7. Figure 6e-h is confusing. It appears as though there is no significant difference in pre- vs post-contact between groups based on Figure 6f-h, yet Figure 6e demonstrates a noticeable difference. Taking 30 seconds pre vs post is likely masking the effect when clearly there seems to be a transient difference immediately after pup contact that is important to capture. Again, it would be helpful to take the AUC at smaller time bins and compare. This would be more reflective of the actual difference.
8. An image of the GCaMP6s transduction into the PVN and example of fiber optic placements would be useful to the reader.
9. Did tethering of the animals cause an overall reduction in movement in any of the tasks compared to untethered animals? Was this quantified?

Minor:
Line73-75: "Likewise" should be replaced with "In contrast" or something to that effect.

Reviewer #3

(Remarks to the Author)
This is a manuscript about an important but understudied phenomenon – how the brain adapts in response to motherhood. While it is known that maternal glucocorticoid responses are blunted during the postpartum period, there have been no direct measures of the central regulator of this stress response, the activity of the corticotropin releasing hormone (CRH) cells within the paraventricular nucleus of the hypothalamus. The authors present several experiments during which CRH activity is measured via fibre photometry in several scenarios, including basal, stress, and in response to pups in both virgin and lactating female mice. Several important distinctions are made about the neuronal activity of CRH neurons in these different scenarios. Overall, this is a well done manuscript that provides important functional information about CRH neuronal activity in females. There are some generally minor, although important, concerns listed below.

The results in 2C are interesting, but seem perhaps to be influenced by one high responder in the virgin group. Was there any outlier analysis done for this (or other) outcomes? If so, what was it? If not, why?

There is a discrepancy in the way pup retrieval results are reported. In the text (lines 204-205), it is stated that there are no differences across groups. However, in the supplement (Figure 1D), there is a main effect of group, with a significant difference between virgin and lactating females in time to pup in nest.

There are some details in the methods that could be expanded upon within the Supplement.
a) How long did it take (on average) for females to become pregnant? The manuscript states they were left until pregnancy was confirmed, but there is no range here.
b) Why was the minimum number normalized to four per litter? What were the maximums and ranges? Where did the extra pups originate? Do the authors think it matters how many pups were in a litter? One assumes so, since care was taken to normalize the minimum.
c) The fibre photometry system should be described in more detail within this manuscript. How much variability was there in the normalization (at least 7 or 3, but up to how many days?)? Given the timing of the postpartum recordings, some animals were pregnant during this habituation period. Do the authors think this extra handling during pregnancy should be considered as readers extrapolate these findings to their own studies?
d) What time within the light cycle (light or dark?) was fibre photometry and behavior recorded?
e) Why PN4 for starting with baseline recordings?
f) How (and when) was tissue collected for RNAScope? How were the intensity thresholds determined for image analysis?
g) Would the authors consider making their custom Python pipeline available? Particularly as they note it is ‘similar’ (but not exactly the same) as those previously described.
h) There are no methods described for blood collection for corticosterone measurement, not how corticosterone is actually measured.

The figure legends (and accompanying text in results) would benefit from being clear about the number of days postpartum during which the recordings were taking place. It is outlined in the methods that there are at least two different cohorts, and that recordings did not always happen on the same postpartum day. An overall schematic figure with the timeline of the different cohorts would also generally improve the ability of the reader to feel confident about what the experimental design was.

Indeed, for the white noise and novel environmental stressors, mice were tested at postpartum days 5 and 6, respectively. For the novel object and pup exposure tasks, mice were tested at postpartum days 7-9, despite these two sets of tasks being performed in different cohorts of mice. The most striking differences in neuron activity were all observed on postpartum days 7-9. The discussion could benefit from a consideration of whether time since parturition plays a role in these findings.

The statistical output is reported very thoroughly in the supplementary materials. This is excellent and appreciated. However, it is reported somewhat inconsistently in the main text.
a) At times, p values are reported in detail (lines 81-97), but this is not the case throughout the manuscript. Other times, they are specified in the figure legends (Figure 2, Figure 3), but this is not the case for every figure legend.
b) At times, the statistical tests used are specified in figure legends (Figure 7), but this is not consistent.
c) There seems to be a typo in the supplementary table – the exact same F/p values are reported for Fig 4 C interaction and Fig 4 D one-way ANOVA

Version 1:

Reviewer comments:

Reviewer #2

(Remarks to the Author)
I have been asked to comment on both the responses to my reviews and that of Reviewer 1, who was not able to re-review the manuscript. I do not share the many of the same concerns as Reviewer 1. I thought the manuscript was quite interesting and novel. Furthermore, while the question of regional heterogeneity is always interesting, it is obvious that fiber photometry

of a small dorsal midline region cannot resolve these questions. Furthermore, this would be challenging with even the most sophisticated 1p and 2p technology given the location of the PVN and may sacrifice the ability to conduct freely moving experiments. Thus, that line of inquiry (Concerns 1-3) is clearly beyond the scope of the study and the authors answered this sufficiently.

For Reviewer 1, point 4 - I do think it is worth mentioning that tethering mice may introduce a delay in the latency to retrieve pups as it generally suppresses movement (I have seen this in my own work). I think this is a caveat that bares repeating considering how widespread fiber photometry has become. Indeed, I make mention of this in my review (Reviewer 2, pt 9). Despite not having quantified this, I this it is nevertheless, worth providing this reasonable caveat in the discussion.

For the rest of Reviewer 1 points 5-9, the authors have answered these comments sufficiently.

My concerns had mostly to do with a) more detailed description of methodologies b) additional statistical analysis and c) addition of data to validate placements. All and all, the investigators have been highly responsive to my concerns and I am happy with them to move forward with publication. Please just add what I mentioned above - but I don't need to see again. Congratulations!

Reviewer #3

(Remarks to the Author)
The authors have provided thoughtful comments and meaningful improvements to the concerns of both reviewers. This manuscript is in excellent shape and will be highly informative to our understanding of how the brain changes during dynamic periods like postpartum.

We would like to thank the three reviewers for their constructive review of our manuscript. Point by point responses to all comments are listed below.

Reviewer #1:

1) The longer I thought about these results the less I was so surprised by them. The authors state starting in the abstract that the lack of suppression of CRH cell activity is contrary to expectation. But just the well known decrease in CRH synthesis by those PVN cells plus the effects of lactation on the pituitary wouldn't further require suppression of CRH cell activity rate to reduce the HPA axis response during lactation. So not entirely a surprise? The increase in activity seen in response to the pup in the metal ball is more notable, but because these cells synthesize and release other neurotransmitters, and these experiments didn't examine what they're releasing in any terminal site, I think there's potential for over-interpretating these results. I also think this result isn't completely surprising in the context of the Walker and Woodside work 20 years ago cited by the authors in the discussion showing that even the peripheral HPA axis response can be strongly activated if the stress is in the presence of the pups.
We appreciate the reviewers thoughts on our experiments.
While prior expectations can certainly guide scientific inquiry, the true function of a brain region can only be understood through careful experimentation. The general consensus in the field is that hypothalamic stress responses are suppressed during lactation. For this reason, we view our work as a valuable contribution, as it provides the first direct characterization of how the real-time activity of hypothalamic stress circuits is modulated during motherhood.

2) There's no consideration of the heterogeneity of PVN CRH cells. Supp Fig 1 seems to show the entire dorsal PVN infected with virus at that level. There's no depiction or description of where the optic implants were located but maybe the recording area is so large and diffuse it doesn't matter. What percentage of the PVN CRH cells in these female mice were infected? For the cells that show fast activity changes to pups, do the data across subjects provide any insight into where they were in the PVN? How are they different from other CRH cells nearby? Does that population project somewhere that's different from other CRH cells that don't respond to pups? Is there any literature on heterogeneity of PVN CRH cells that could help the data interpretation? I think the manuscript would be strengthened if some of these or other questions were answered. Expanding the discussion despite the journal's word limit would also help.

Below we address each of the reviewers points.
CRH heterogeneity, optic fibre implant and expression level: While there is heterogeneity within the CRH population, this heterogeneity is not well understood in the field. We do however know that in this mouse line, the majority of CRH neurons in the PVN project to the median eminence (Wamsteeker Cusulin et al 2013). In addition, fibre photometry uses a large diameter optic fibre which collects light from much of the PVN meaning that differentiating between different subregions based on fibre placement is not feasible. We have now added a line to the methods stating

where we aimed the location of the tip of the optic fibre in the PVN. We have previously validated GCaMP expression using AAV injection into the PVN and found that 58.1±2.1% of CRH neurons on one side of the PVN were transduced (Kim et al, 2019). This current study has used the same techniques and the virus injections were performed by the same researcher. For validation in the previous study, we used a CRH-cre mouse crossed with a tdTomato reporter line to visualise all CRH neurons. For this current study, we did not cross mice onto this reporter line and so do not have a good way to label all CRH neurons. Based on experience from our past studies, we set a minimum threshold of 10% ΔF/F photometry response to pick up stress with animals exhibiting smaller responses being excluded. This ensured we only recorded from mice with good expression and fibre placement.

For the cells that show fast activity changes to pups, do the data across subjects provide any insight into where they were in the PVN? How are they different from other CRH cells nearby? Does that population project somewhere that's different from other CRH cells that don't respond to pups? Is there any literature on heterogeneity of PVN CRH cells that could help the data interpretation? I think the manuscript would be strengthened if some of these or other questions were answered. Expanding the discussion despite the journal's word limit would also help.

There is limited literature investigating projection-defined CRH neuron populations. To our knowledge, the soma of CRH neurons that have different central brain projections do not occupy clearly distinct and demarcated regions of PVN. Instead, these neurons are intermixed. As noted above, due to the fact that the optic fibre collects emitted fluorescence from a large area of the PVN, it would also not be possible to clearly distinguish subregion specific responses.

In the revised manuscript, we have added additional discussion of this.

3) I'm wondering whether or not 24 hours separated from pups was long enough to restore the cell activity function of the PVN CRH cells as detectable by the photometry. They cite studies showing that 24 hours is enough to restore the CRH mRNA synthesis in those cells, but could a longer separation have been better for this study? Related to separation in the novel object experiment, the lactating mice have pups taken away I think for a range of 2-4 days before recording. Did this range create heterogeneity in the results? Is there more to dig into the data there? I also wonder why mice remaining with their pups were removed from them 1 hour before recording and whether or not this affected the results.

Past work that we cited has shown that the corticosterone response in lactation is suppressed with pups and remains suppressed if pups are removed for one hour. However, if pups are removed for 24 hours, the corticosterone response goes back to control levels (Kalyani et al, 2017). This study also compared the corticosterone stress response if dams were separated from pups for 24 hours vs 8 days and found no significant difference. This suggests that increasing the duration of pup separation will not have a large impact.

As the reviewer notes, for the novel object and pup interaction tests, dams were separated from pups for 1 hour prior to the test. Removal of pups involved an experimenter reaching into the cage, which is mildly stressful for the dam. 1 hour was also the time given for mice to settle (and reach a stable baseline) after this necessary manipulation. The study above suggests that 1 hour separation should

not have an impact on the normal suppression of stress responses seen during lactation.

4) The latencies to contact pups is very long in Supp Fig 1, at about three minutes, and makes me concerned about the behavior of these animals and whether they're somehow stressed at baseline and not responding properly. Was the lack of group difference in that and the other behaviors expected when it was the very first time the virgins were shown the pups?
A baseline recording of at least 30 minutes was taken to ensure activity was stable and low. As such, we would be able to see if neural activity was markedly elevated indicating animals were stressed. We did not observe this. However, we do appreciate that having mice tethered to an optic fibre may interfere with normal behaviour. This is the most likely explanation for the slow initial contact with pups.

Was the lack of group difference in that and the other behaviors expected when it was the very first time the virgins were shown the pups?
Previous work (Swart et al., Hormones and Behavior, 2021) has shown that there is no significant difference between pup naïve virgin mice and lactating mice in the latency to approach the first pup in a pup retrieval assay. However, lactating mice have a significantly shorter retrieval latency for the first pup. These findings are the same as what we have observed.

5) Could baseline firing of PVN CRH cells differ between virgins and lactating mice and affect the photometry data when the data for each group is relative to different baselines? Maybe this is accounted for in the data analysis in ways I'm not clear about.
We have carefully determined if baseline activity was different between virgin and lactating mice (Figure 1). Rather than average baseline activity, we analysed the amplitude, frequency and slope of discrete photometry events that make up baseline activity. We found no significant difference between virgin and lactating groups. Importantly, our previously published research has demonstrated that this analysis approach can detect significant changes in baseline activity in the CRH neuron population. Specifically, we could detect changes in baseline activity induced by corticosterone negative feedback (Kim et al, 2019).

6) The authors may want to emphasize in the discussion that it's only the PVN CRH neurons they targeted so the results may not extend to how other groups of CRH cells function. Is anything known about other CRH cells in other sites during lactation?
We have now clarified that our results only pertain to PVN CRH neurons. We are not aware of any other studies which have looked at the activity patterns of other CRH neuron populations during pregnancy/lactation.

7) The authors can specify what CRH cell population they're referring to in each instance. I didn't know when they were talking about their or other studies whether or not they always meant it was the PVN population.

In all cases, we refer to PVN CRH neurons. We have now added additional clarification in the text.

8) Is the X axis on Supp Figure 2 the total number of pup calls? Can trend lines be added to those panels to help make the group differences in the correlations more obvious? I don't know what to make of the data in panel G because that one dataset doesn't have the full distribution of calls compared to the other two. Is that a statistical concern when comparing correlations between groups when the underlying data distributions differ? Maybe convert data to Z-scores? The Y axis on that panel also differs from the others for some reason.

We believe that the distribution of the data shown in panel G is quite similar to that in panel E with the only difference being the small number of data point with high USV calls in panel E. We have now made corrections to these figures as suggested by the reviewer.

9) Animals paragraph in methods - after pregnancy was confirmed, males were removed but were the females housed singly or group housed? The minimum number of pups was 4 but what was the maximum and do the data from the study show whether the number of pups being raised affects the maternal CRH cells or their behavior? If the range of litter size is large, maybe that needs to be a covariate.

The maximum number of pups was limited to 8. Overall, only one dam had 4 pups. The average pups per lactating dam was 7 ±0.4 pups. This is now reported in the manuscript.

Reviewer #2:

1. The description of the statistics used are insufficient. There are several cases in which use of two-way ANOVAs (in some cases repeated measures, in some cases not) are required in order for the investigators to make their conclusions. Yet it is unclear that they are using the appropriate statistics: Figure 2c,3c,4c, 5d,f,g,h, most of the comparisons in Figure 6. The only place this is indicated is in the Figure legend of Figure 7. For within subject comparisons, the investigators can normalize to baseline and then run one-way ANOVAs. It may also help the reader to show the normalized responses in the main text and raw data in an extended figure for Figure 6 since most of the comparisons appear to be within subject baseline vs. post-stimulus. In some places this is done and in others it is not. It would be good to be consistent in how the data are displayed across figures.

Full descriptions of statistical analyses are provided in supplementary table. We now also provide statements of the key statistical results in the main text. All data were normalised to the group mean baseline. This has been added to the methods (line 432-433).

2. Provide technical details of fiber photometry recordings in the methods rather than just referencing the two citations. Specifically, was use of an isosbestic signal (e.g. 405) employed. If not, how were issues like motion artifact corrected for?

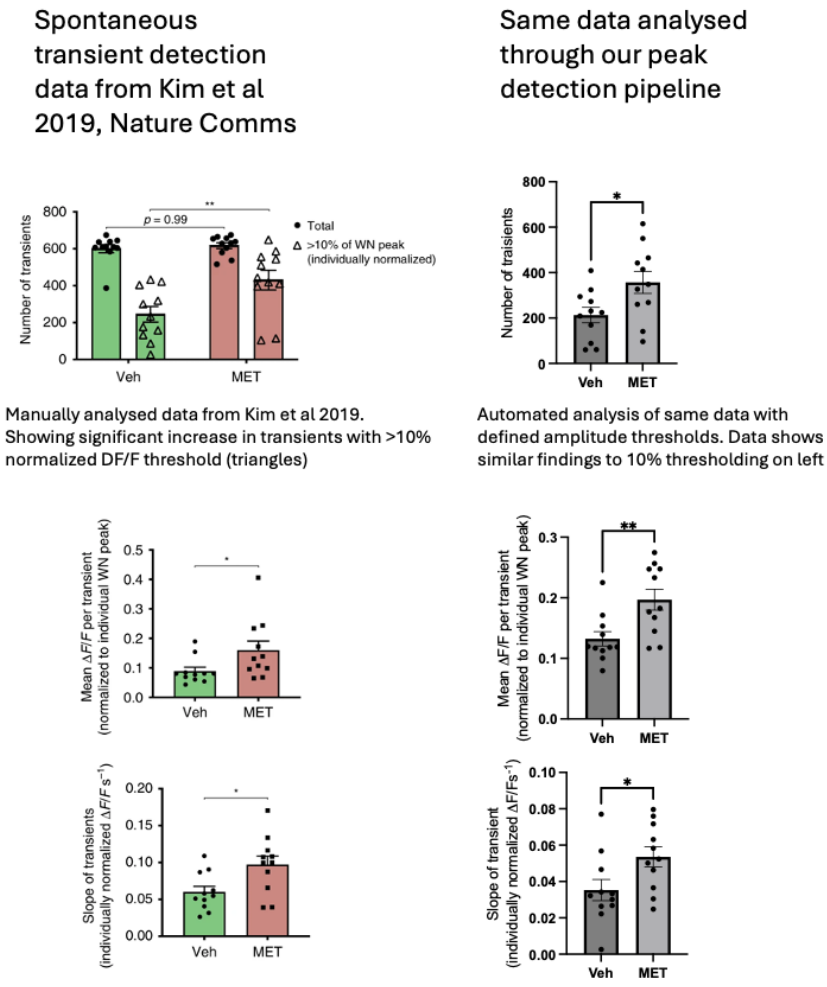
We apologise for not providing these methods previously. They are now added to the data analysis section of the methods.

Motion artefacts were subtracted by fitting the isosbestic signal to the 465 signal. Overall, we had extremely little motion artefacts, largely due to our selection criteria for high signal recordings. Our mean peak $\Delta F/F$ signal across all mice was 53.8% ±0.37. Such recordings were minimally affected by noise.

3. Normalization of event-related transients is standard; however, normalization of spontaneous transients is problematic. It is better to report the raw frequency and peak amplitude. See Wallace et al., 2025, ACS Chemical Neuroscience.

The reviewer's concern is valid and shared by us. Wallace et al report that Z-scoring normalisation is problematic and we note that our data is not Z-scored. Instead, our data was normalised to the peak $\Delta F/F$ for each subject. We applied 3 filters to the scipy find_peaks function. 1) A raw $\Delta F/F$ threshold of 3%, applied to all animals. 2) A 5% normalised $\Delta F/F$ threshold (this resembles the animal-specific cutoff noted by Wallace et al., 2025). 3) Since GCaMP6s is slow, a temporal threshold of >0.5s was also applied to the trough to peak window, which removed fast artefact noises. We have added this to the methods (from line 455).

Prior to analysis, we tested this method on prior published manual peak detection data from Kim et al., 2019 Nature Communications [PMID: 31836701]. Below are data from Fig 3J-I from the paper (left), shown in both manual detection and validated by our automated detection (right). This is presented for reviewer purposes only.



Reviewer Figure 1. Comparison of manual detection vs automated detection of previously published data.

4. In addition to reporting the cumulative signal (and also AUC would be useful) for the entire 5 or 10 minute duration of the stimulus in Figure 2 and 3, the investigators should isolate and expand the time window of the onset, sustained and offset components. It may be that that by not isolating and analyzing different components of the stimulus presentation, they are missing valuable information regarding how PVNCRH neurons are differentially encoding environmental stressors in virgin vs. lactating females.

It is tempting to pick out small time windows for detailed analysis, however, the risk here is that this increases our chances of type 1 error. Instead, we have tried to take a less biased approach which looks for differences across the entire recording. Our cumulative integrated ΔF/F is essentially plotting a running AUC over all time points of the stress response from the very beginning to the end. We also plot the post-stress response separately and show a notable difference between groups here. In summary, we believe our analysis captures the key (biologically meaningful) differences between our groups where they exist.

5. The methods for collecting corticosterone are absent. It is unclear if the "basal" and "whitenoise" blood is collected within group or in separate sets of mice. If it is the latter, there should not be connecting lines. Seems like it is within subject…in which case, is it tail blood that is being collected?
We apologise for this oversight. All bloods were obtained by tail sampling and the data shown in Fig 4 are repeated samples (hence the connecting lines). We have now added description of the blood sampling to the methods (line 429)

6. Clarify that what is measured is "area/cell" in the y-axis labeling in 4b.
This is noted in the methods as the area of the mRNA labelling as a percentage of the defined total PVN area (line 450)

7. Figure 6e-h is confusing. It appears as though there is no significant difference in pre- vs post-contact between groups based on Figure 6f-h, yet Figure 6e demonstrates a noticeable difference. Taking 30 seconds pre vs post is likely masking the effect when clearly there seems to be a transient difference immediately after pup contact that is important to capture. Again, it would be helpful to take the AUC at smaller time bins and compare. This would be more reflective of the actual difference.
We agree that 30s bins can mask fast changes, particularly for data describing acute contact. Figure 6 E indeed suggests increased activity in lactating and pups removed groups and if we reanalyse this data with 10s bin comparison, we find a significant difference (Effect of lactation ($F(2, 16) = 0.215$, $P=0.80$); Effect of time ($F(1,16)=7.27$, $P = 0.0001$); Interaction ($F(2,16)=1.53$, $P=0.23$). Post hoc test results: Virgin pre vs post, $P= 0.44$; Lactating pre vs post, $P=0.019$; Pups removed pre vs post, $P = 0.003$. We have now reported this in the results text.

8. An image of the GCaMP6s transduction into the PVN and example of fiber optic placements would be useful to the reader.
An image of GCaMP6s transduction is included in Supplementary Figure 1.

9. Did tethering of the animals cause an overall reduction in movement in any of the tasks compared to untethered animals? Was this quantified?
This is likely, however we are not able to compare directly with non-tethered animals for this project. It is worth noting the majority of our experiments are in the home cage, in which mice are typically less active. While it is known that tethered animals show some movement perturbations, this is a limitation we are unable to overcome for this project.

Minor:
Line73-75: "Likewise" should be replaced with "In contrast" or something to that effect.
Modifications to the text have been made based on the reviewers suggestion.

Reviewer #3:
1) The results in 2C are interesting, but seem perhaps to be influenced by one

high responder in the virgin group. Was there any outlier analysis done for this (or other) outcomes? If so, what was it? If not, why?
No outliers were detected (ROUT analysis). We also note that removal of the one high responder in virgin group does not alter the results.

2) There is a discrepancy in the way pup retrieval results are reported. In the text (lines 204-205), it is stated that there are no differences across groups. However, in the supplement (Figure 1D), there is a main effect of group, with a significant difference between virgin and lactating females in time to pup in nest.
Thank you for spotting this. The text has now been corrected.

3) There are some details in the methods that could be expanded upon within the Supplement.
a) How long did it take (on average) for females to become pregnant? The manuscript states they were left until pregnancy was confirmed, but there is no range here.
Unfortunately we did not obtain reliable vaginal plugs for all animals. Pregnancy was confirmed visually and by weight increase, which was evident during mid-gestation. Overall, pregnancies were confirmed within 3 weeks from the time a male mouse was introduced. From this, we can infer pregnancy within the first week. However we do not have the data for the exact days to pregnancy.

b) Why was the minimum number normalized to four per litter? What were the maximums and ranges? Where did the extra pups originate? Do the authors think it matters how many pups were in a litter? One assumes so, since care was taken to normalize the minimum.
Pup numbers were normalised as suckling intensity and number of pups can influence maternal hormone milieu. Maximum litter size was 8 pups, with additional pups cross-fostered. These additional details have now been added to the manuscript (added in line 381).

c) The fibre photometry system should be described in more detail within this manuscript. How much variability was there in the normalization (at least 7 or 3, but up to how many days?)? Given the timing of the postpartum recordings, some animals were pregnant during this habituation period. Do the authors think this extra handling during pregnancy should be considered as readers extrapolate these findings to their own studies?
We have now added additional details regarding the fibre photometry to the methods.
We calculated the std error across all normalisation values within subjects. Overall, the variability of the peak response within subjects was 7.38% $\Delta F/F$. This represents the variability in the peak response across all trials.
Given the amount of handling and tethering habituation employed in our lab, we believe the handling during late pregnancy and early postpartum period had minimal impact. We were careful not to begin handling during sensitive periods of early pregnancy and immediately following birth.

d) What time within the light cycle (light or dark?) was fibre photometry and behavior recorded?
All recordings and behaviour were measured during ZT2-10. This is now reported in the manuscript (line 385).

e) Why PN4 for starting with baseline recordings?
This timeframe was chosen so to avoid disrupting maternal behaviour immediately following birth which could lead to reduced pup survival. This is now added in line 399.

f) How (and when) was tissue collected for RNAScope? How were the intensity thresholds determined for image analysis?
Postpartum day 7 is now noted in results section. Thresholds was set to 50 within the Qupath workflow and is reported in the manuscript. This threshold setting was empirically tested and confirmed to efficiently detect mRNA labelled soma without detecting background labelling.

g) Would the authors consider making their custom Python pipeline available? Particularly as they note it is 'similar' (but not exactly the same) as those previously described.
We now provide a weblink in the manuscript to the code.

h) There are no methods described for blood collection for corticosterone measurement, not how corticosterone is actually measured.
We apologise for this oversight. We have now added description of the blood sampling to the methods.

4) The figure legends (and accompanying text in results) would benefit from being clear about the number of days postpartum during which the recordings were taking place. It is outlined in the methods that there are at least two different cohorts, and that recordings did not always happen on the same postpartum day. An overall schematic figure with the timeline of the different cohorts would also generally improve the ability of the reader to feel confident about what the experimental design was.
We have now added additional text clarifying the postpartum day when experiments were conducted.

5) Indeed, for the white noise and novel environmental stressors, mice were tested at postpartum days 5 and 6, respectively. For the novel object and pup exposure tasks, mice were tested at postpartum days 7-9, despite these two sets of tasks being performed in different cohorts of mice. The most striking differences in neuron activity were all observed on postpartum days 7-9. The discussion could benefit from a consideration of whether time since parturition plays a role in these findings.
We apologise for not correctly reporting the different cohorts of mice. The cohort of mice that was tested with novel object and pup exposure had previously been tested on postpartum 4-6 with basal activity, white noise and novel environment. This is now correctly described in the methods section.

We have also added additional text to the discussion elaborating on this point.

6) The statistical output is reported very thoroughly in the supplementary materials. This is excellent and appreciated. However, it is reported somewhat inconsistently in the main text. a) At times, p values are reported in detail (lines 81-97), but this is not the case throughout the manuscript. Other times, they are specified in the figure legends (Figure 2, Figure 3), but this is not the case for every figure legend. b) At times, the statistical tests used are specified in figure legends (Figure 7), but this is not consistent.
Thank you for pointing out this issue. We have now consistently reported the statistical results in the results manuscript text.

c) There seems to be a typo in the supplementary table – the exact same F/p values are reported for Fig 4 C interaction and Fig 4 D one-way ANOVA
Data in Figure 4C was analysed with a two-way ANOVA whereas data in Figure 4D was analysed with a one-way ANOVA. As such, the values reported in the table are correct.