

## **Distribution Agreement**

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Rahil Mahmood

April 03, 2023

A Machine Learning Network to Automatically Track Prairie Voles in Cohabitation:  
Oxytocin Receptor KO Males Reveal No Behavioral Deficits Towards their Partners

by

Rahil Mahmood

Dr. Robert Liu  
Advisor

Biology

Dr. Robert Liu  
Advisor

Dr. Kate O'Toole  
Committee Member

Dr. Mar Sanchez  
Committee Member

2023

A Machine Learning Network to Automatically Track Prairie Voles in Cohabitation:  
Oxytocin Receptor KO Males Reveal No Behavioral Deficits Towards their Partners

By

Rahil Mahmood

Dr. Robert Liu

Advisor

An abstract of  
a thesis submitted to the Faculty of Emory College of Arts and Sciences  
of Emory University in partial fulfillment  
of the requirements of the degree of  
Bachelor of Science with Honors

Biology

2023

## Abstract

### A Machine Learning Network to Automatically Track Prairie Voles in Cohabitation: Oxytocin Receptor KO Males Reveal No Behavioral Deficits Towards their Partners

By Rahil Mahmood

Quantifying the nature of social interactions displayed by prairie voles in cohabitation can be a useful way to understand the neural mechanisms underlying various social behaviors. However, the current standard of behavioral analysis involves the annotation or scoring of experimental recordings, which in addition to being a time-consuming process, is open to biases and great variability across human annotators. Using supervised machine learning principles, we developed a robust network capable of automatically tracking prairie voles in cohabitation. Trained with the help of the open-source pose estimation tracking software DeepLabCut, our tracking network incorporated an in-house autoencoder and was built on a dataset of over 500,000 frames of video. Ultimately, the network was capable of increasing the efficiency of the annotation process by nearly 100%. Oxytocin has previously been shown to play a crucial role in regulating the social behaviors that govern the formation of pair bonds in prairie voles and other monogamous mammals. We used our auto-tracking network to quantitatively compare the differences in social behaviors displayed by wildtype and CRISPR/Cas-9 mediated oxytocin receptor knockout male prairie voles in cohabitation with their wildtype female conspecifics. We found no significant differences between the two genotypes with regards to the average total duration and frequencies of prosocial behaviors displayed at the time of cohabitation. These results were consistent across both automated and manual tracking processes. In concordance with previous studies, we posit that pair bond maintenance, and hence the maintenance of the social behaviors regulating pair bonds, can be achieved in the absence of the oxytocin receptor. This is likely due to the cross-talk that has been shown to exist among the oxytocin, vasopressin,

and dopaminergic pathways in the brain. Ultimately, we propose that our automated tracking model can be used to significantly eliminate the subjectivity inherent in the process of behavioral annotation and eventually serve to facilitate the standardization of annotation protocols in the field of computational neuroethology.

A Machine Learning Network to Automatically Track Prairie Voles in Cohabitation:  
Oxytocin Receptor KO Males Reveal No Behavioral Deficits Towards their Partners

By

Rahil Mahmood

Dr. Robert Liu

Advisor

A thesis submitted to the Faculty of Emory College of Arts and Sciences  
of Emory University in partial fulfillment  
of the requirements of the degree of  
Bachelor of Science with Honors

Biology

2023

## Acknowledgements

I would like to thank Dr. Robert Liu for accepting me into his lab and supporting me over three years on various projects, including this one, and giving me invaluable guidance on carrying out research as an undergraduate. I am grateful for all the opportunities that he has presented me with to carry out independent assignments, share my findings at lab meetings, and most importantly, gain exposure to the rigors of professional research.

I would also like to thank Dr. Hong Zhu, who has directly mentored me over my three years in the Liu Lab, hand-holding me through all my projects, teaching me various practical and theoretical skills, and giving me access to all her data to work with—including on this project. Dr. Zhu has played a monumental role in helping me put this project together and I am extremely grateful for all the time she has dedicated towards ensuring my success. Her mentorship has been invaluable and I will forever be indebted to her for providing me with the resources to investigate this topic.

I would also like to thank Dr. Sena Agezo for giving me access to his automated tracking pipeline, teaching me how to use it, and inspiring me to pursue this topic. I could not have undertaken this project without his mentorship, granting of access to his data, and his invaluable insights on perfecting the tracking model. This project is essentially a subset of a larger project previously pursued by Dr. Agezo.

I would also like to thank Jenny Zha for assisting me with data collection. I could not have completed this project if not for her sincere efforts and countless hours spent in improving the results of the automated tracking process.

Finally, I would like to thank all members of the Liu Lab who supported me at various stages during my undergraduate research experience, provided me with opportunities to present my thoughts at lab meetings, and inspired me to keep pushing forward when times were tough.

## Table of Contents

<b>INTRODUCTION</b>	1
<b>METHODS</b>	7
Animals	7
Behavioral Recording and Manual Annotation	7
Training Neural Network (DLC)	8
Autoencoder	9
Data Analysis	11
<b>RESULTS</b>	17
Comparing Auto-Tracked and Manually-Derived Results	
Independent of Genotype	17
Mating	17
Social Investigation	18
Comparing Quantitative Differences in Social Behavior	
Across Genotypes	19
<b>DISCUSSION</b>	20
Future Directions	26
<b>REFERENCES</b>	28
<b>APPENDIX</b>	32
MATLAB Code	32



## FIGURES

1. Points used to label vole via DeepLabCut (DLC)	9
2. Determining instances of nose-nose sniffing	13
3. Determining instances of anogenital sniffing	15
4. Determining instances of side sniffing	16
5. Determining instances of mating	16
6. Correlation between auto-tracked and manually generated results (mating duration)	45
7. Correlation between auto-tracked and manually generated results (social investigation duration)	46
8. Correlation between auto-tracked and manually generated results (mating instances)	47
9. Correlation between auto-tracked and manually generated results (social investigation instances)	48
10. Distribution of average durations of social behaviors (auto-tracking)	49
11. Distribution of average durations of social behaviors (auto-tracking)	50
12. Distribution of average durations of social behaviors (manual-tracking)	51
13. Distribution of average number of instances of social behaviors (auto-tracking)	52
14. Distribution of average number of instances of social behaviors (auto-tracking)	52
15. Distribution of average number of instances of social behaviors (manual-tracking)	53
16. Correlation between auto-tracked and manually generated results (social investigation durations)	54
17. Cumulative probability distribution of frames identified by auto-tracking network for various male-female distance thresholds	55

18. Distribution of $R^2$ and slope values obtained for correlations between manual and auto-tracking results across all animals (n=11)	56
19. Frames identified by auto-tracking network for instances of nose-nose sniffing at a specific threshold value	57

## Introduction

The neural mechanisms underlying social behavior in mammals have long been of interest in the field of behavioral neuroscience. Pair bonding, or the establishment of a long-term emotional connection facilitated through exclusivity in mating, or social monogamy between conspecifics (Bales et al, 2021), is observed in merely five percent of mammals, including prairie voles (*Microtus ochrogaster*) and humans (Johnson and Young, 2015). Pair bonding behavior is believed to be evolutionarily advantageous (Fletcher et al., 2015) in certain environments by enhancing reproductive performance, promoting biparental care, and increasing cooperation between members in harsh environmental conditions (Sánchez-Macouzet et al., 2014). Among some of the key hormones and neuromodulators responsible for regulating the social behaviors associated with establishing and maintaining pair bonds, oxytocin has been shown to be especially important (Burkett et al., 2016) due to its effects exerted across multiple behavioral-regulating regions of the brain, including the nucleus accumbens (NAcc), prelimbic cortex (PLC) in the medial prefrontal cortex (mPFC), and the bed nucleus stria terminalis (BNST), to name a few (Horie et al., 2020; Liu and Wang, 2003). While previous studies have investigated how the global knockout of oxytocin receptors (Oxtr-KO) in male prairie voles does not prevent them from maintaining a preference towards their WT female partners (Horie et al., 2020), a shortage of data and lack of clarity exist in terms of understanding the specific roles of oxytocin receptors in regulating the nature of affiliative and non-affiliative behaviors that individuals display when encountering familiar faces. Specifically, there is a need for more clarity on how oxytocin receptors help modulate the prosocial behaviors (mating, sniffing, huddling, etc.) that male prairie voles often display when placed alongside their female partners. While the use of partner preference tests has previously served as a useful experimental set-up to investigate the role of

oxytocin receptors in maintaining preferences in male prairie voles for partner females over novel females (Beery, 2021; Williams et al., 1994), in order to look for the specific types of social and non-social behaviors that male prairie voles may display around their female partners, it is more useful to have the animals in cohabitation (López-Gutiérrez et al., 2021). In this setting, both the male and female are untethered and free to move around the entirety of the experimental environment (unlike in a partner preference test), thereby facilitating the enactment of behaviors in their most natural form (i.e. as would be seen in the wild).

To modulate the oxytocin system, some previous studies infused an oxytocin receptor (OTR) antagonist into the NAcc and PFC and found that doing so restricted the display of partner preference in female prairie voles (Young and Wang, 2004). Other studies found that in male prairie voles, the infusion of an OTR antagonist into these same regions could prevent males from maintaining partner preference (Wang and Aragona, 2004). Given that there exists significant crosstalk between the oxytocin, arginine vasopressin (AVP), and dopamine receptor systems in these cortical regions (Song and Albers, 2018; Baribeau and Anagnostou, 2015; Rae et al., 2021), and that there are inefficiencies associated with antagonist usage (Josipović et al., 2019), the aforementioned studies have been limited in their ability to fully understand the role of the oxytocin receptor in regulating social behaviors and social preferences.

With the advent of modern biotechnologies in recent years, it has been possible to overcome many of the challenges and uncertainties posed by the use of receptor antagonists. One such technology is CRISPR (Heidenreich and Zhang, 2016). The CRISPR/Cas-9 facilitated oxytocin receptor knockout line of prairie voles has previously served as a useful model for demonstrating that the global knockout of oxytocin receptors leads to repetitive behaviors and

impaired preferences for social novelty (Horie et al., 2019). Some recent studies have used the CRISPR/Cas-9 model to demonstrate that *Oxtr*-KO prairie voles retain the ability to form social attachments and also maintain them over an extended period of time (Berendzen et al., 2023). However, studies like these primarily focused more on quantifying whether or not bond formation took place and less on *how* it took place. In other words, the question on how exactly the impairment of oxytocin signaling influences the numerous prosocial behaviors that prairie voles display *during* pair bond formation (or when they first encounter a new face) remains unexplored. Additionally, as insightful as these studies have been, many of the standard experimental set-ups that they have incorporated to objectively measure the deficits in social behavior seen in *Oxtr*-KO individuals are constrained by the requirement of the human element (i.e., the need for human scorers to manually annotate specific instances of different social behaviors by viewing video recordings of prairie voles in cohabitation) (Noldus et al., 2001).

While having human annotators may be advantageous due to their ability to observe and differentiate between certain behavioral nuances (a skill that comes from experience and familiarity with prairie vole behavioral patterns), the process that they must partake in to annotate videos is often quite time consuming (Silverman et al., 2010; Winslow, 2003). Moreover, the process is vulnerable to significant bias and variability when different annotators work on different videos for the same experiment (Shemesh et al., 2013). Furthermore, this process of manual annotation does not allow for the development of a fixed standard of behavioral definitions that can be adopted and shared across labs interested in investigating similar questions (Segalin et al., 2021). For instance, determining whether a male “sniffs” or investigates a female, or vice-versa, is dependent upon the annotator's definition for the distance threshold between the two animals that would be appropriate to facilitate the act of “sniffing”. If

this threshold is chosen to be  $x$  units of distance, then the annotator would be required to visually judge this distance as accurately as possible on every occasion when the animals are in proximity of each other (and often moving at high speeds). Doing so is often beyond the scope of human sensory ability, and the prospect of training annotators to acquire the same reaction time is unfeasible—even when all the annotators work in the same lab. Inevitably, when the same behavior is assessed under different experimental conditions at different labs, the absence of a commonly agreed upon definition for investigative threshold distances or threshold mating angles, for example, could lead to a lack of generalizability in the results generated across different labs.

To combat the many flaws inherent in the process of manually annotating social behaviors, there have been several efforts in the field of computational neuroscience to create more “automated” approaches—capable of addressing the previously mentioned time, bias, and standardization factors (Sturman et al., 2020). These approaches have primarily fallen under two broad categories of machine learning: supervised and unsupervised. The former derives its name from the fact that it is trained on human-generated “ground truth” data. In other words, human annotators must create a data set consisting of correctly-labeled samples that the automated network can then use as a reference point. The latter, meanwhile, requires no human input or ground truth. Instead, it relies on Hidden Markov Models to look for repetitive patterns in the data and then classifies different behavioral sequences according to these distinct patterns (Wiltschko et al., 2015). In the case of both the supervised and unsupervised machine learning approaches, there are certain drawbacks to be considered. The requirement for ground truth in the case of the supervised approach makes it open to biases that might be present in the human-

generated data. Meanwhile, in the case of the unsupervised approach, the sequences recognized often tend to be overly complex (representing more behaviors than there actually are), and there is therefore a requirement for human subjects to ultimately deconstruct these complex patterns. Here too, then, there is certainly room for human bias, depending on how the data are interpreted. However, unlike in the case of manual annotation, the “bias” in the case of both supervised and unsupervised auto-tracking approaches has nothing to do with human reaction times or visual sensory capabilities. Rather, it arises mostly at the time of post-tracking analysis and its effects are often less detrimental than those that result from having badly tracked data to start with. Therefore, with both supervised and unsupervised approaches offering a significant reduction in the time factor and annotation bias—when compared to manual annotation processes—the decision regarding which approach to choose boils down to the question of generalizability in the results. As different labs may be interested in different variations of certain social behaviors (for example, general “sniffing” behaviors vs specifically “oral sniffing”), many previous studies (Nilsson et al., 2020; Walter and Couzin, 2021; Pereira et al., 2020) have leaned towards supervised tracking approaches due to the flexibility they offer in allowing researchers to predefine specific behaviors of interest in the ground truth. However, the subject of generalizability and the extent to which is it possible, is still heavily debated (Segalin et al., 2021).

With the goal of attempting to standardize the behavioral annotation process within and across labs, drastically reducing the time factor of the manual labeling process, and effectively eliminating the bias and variability inherent in manual annotation, this study incorporated an automated tracking methodology capable of increasing the efficiency of the tracking process by nearly one hundred percent. The open source, pose estimation tracking network DeepLabCut

(DLC) (Mathis et al., 2018) was trained with a large sample of manually-labeled tracking data and a multi-layered autoencoder was developed in-house to filter out the inaccuracies in the DLC tracking estimations. This robust, supervised machine learning network was then used to evaluate the differences in the nature of social interactions between five wild type (WT) and six oxytocin-receptor knockout (OXTR-KO) male prairie voles with their WT female partners during cohabitation. While some previous studies (Kitano et al., 2022; Berendzen et al., 2022; Ross et al., 2009) have looked into similar interactions, their analyses have relied primarily on manual annotation processes—thereby casting doubts on the generalizability of such findings as outlined above. Our study, however, is novel in that it comprises an unprecedented sample size of over 500,000 frames of cohabitation video to draw conclusions based on results generated by a highly regulated automated tracking network, and then compares these results to those obtained from conventional manual annotation practices. We found a strong correlation between the results generated by both the manual and automatic tracking approaches, allowing us to share our predictions on the role of the oxytocin receptor in mediating pair bonding behaviors with great certainty. Most of all, our model offers the opportunity to derive conclusions regarding the nature of a particular cohabitation in a time-span that is nearly less than half of that required when adopting the conventional process of manual annotation. In the sections that follow, we describe the steps we took to develop the tracking network, including how we ensured its robustness and reliability, and how we used it to explore a pressing question in the field of behavioral neuroscience. We believe that the significance of our work, however, lies less in its implications on oxytocin and pair bonding, but more in the field of automated behavioral tracking. We strove to define different social behaviors with the utmost clarity, using our familiarity with prairie vole



behavioral patterns to account for all possible physical manifestations—obvious and subtle—of the behaviors we investigated.

## Methods

### 1. Animals

Homozygous oxytocin receptor knockout male prairie voles generated by CRISPR/Cas9 editing (as outlined by Horie et al. (2020)) and wildtype male controls from the same breeding colonies were cohabitated with wild-type females obtained from the stock colony housed in the lab of Dr. Larry J. Young at Emory University in Atlanta, Georgia.

### 2. Behavioral Recording and Manual Annotation

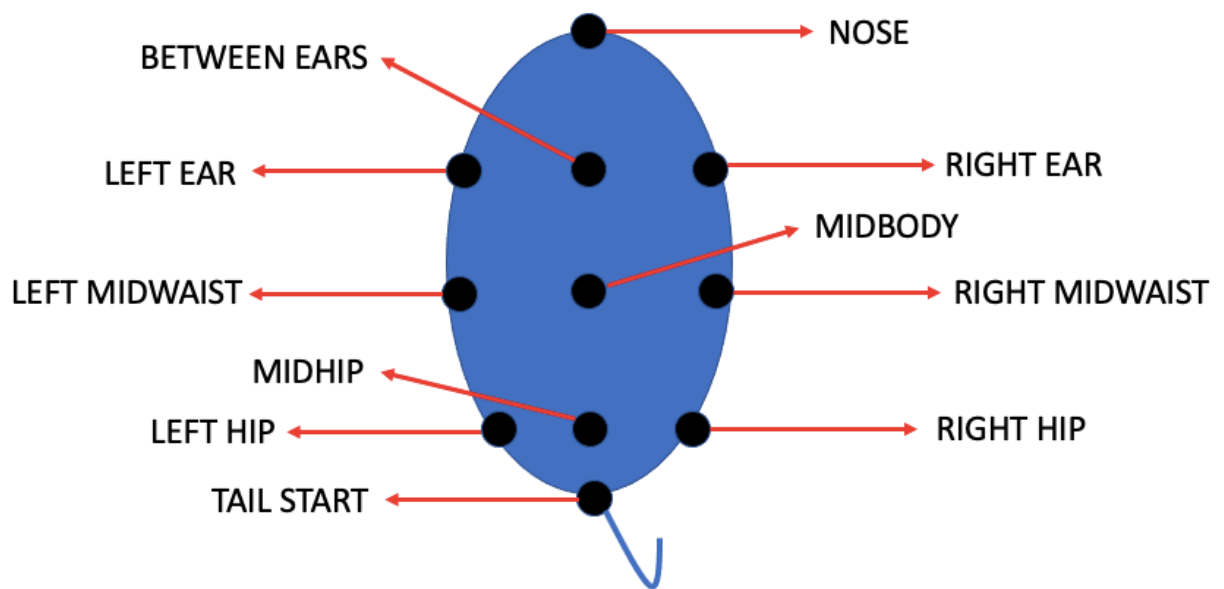
A downward-facing camera was placed above each chamber to record the animals engaging in their native behaviors during cohabitation. The camera frame rate was set to 20 fps (or 0.05 seconds per frame). The open-source event logging software called Behavioral Observation Research Interactive Software (BORIS) was used to create an ethogram of various social behaviors, including *social investigation*, *mating*, and *approach/departure* instances. This ethogram was subsequently used to annotate videos of cohabitation that lasted on average for a duration of 65 minutes. The human annotators were unaware of the genotype (WT/KO) of the animals in the videos, eliminating any potential biases from preconceived expectations on behavioral outcomes. *Social investigation* was defined as any attempt by the male to “sniff” the female within a visually determinable distance of approximately half the length of the average prairie

vole head. Instances wherein the male approached the female head-on (nose-nose sniffing), from either side of the female's body (nose-to-ear/waist/hip), and from the female's rear end (anogenital sniffing) were all included broadly under *social investigation*. The behavior *mating* was defined as any instance wherein the male mounted the female with the intention of intromission. Instances wherein intromission did not occur were classified separately as *mounting only*. Instances of *approach* and *departure* were defined as points wherein the male accelerated towards the female from rest with the intention of investigating it, and when the male accelerated away from the female after investigating/mating with it, respectively. Following the completion of the annotation process, BORIS provided the tools to generate metrics with regards to the total durations of each behavior and an overall timeline of behavioral events throughout the video. Comparing such metrics across different animals (such as total *social investigation* duration in WT vs KO) allowed for inferences pertaining to the effects of OXTR-KO on regulation of pro-social/pair bonding behaviors.

### 3. Training Neural Network (DLC)

The open-source pose estimation package DeepLabCut (DLC), developed by the Mathis Group and the Mathis Lab, was used to create separate 2D skeletal frameworks for the male and female inside each recording chamber (Figure 1). Each skeleton consisted of eleven points that served to provide an outline of the body shape: *nose, left ear, between ears, right ear, right-mid-waist, midbody, left-mid-waist, left hip, midhip, right hip, tail start*. Approximately 50000 frames were labeled per video, ultimately giving rise to a sizable “ground truth” of manually-labeled data (over 500000 frames) to

train the DLC network. Each of the eleven points labeled on the “skeleton” represented an X-Y coordinate for that particular anatomical feature. Each recording chamber was also mapped out to define the coordinates of the four corners and the center. Ultimately, the skeletal coordinates generated by the network were in relation to the coordinates of the cage (with the center of the chamber labeled as 0,0).



**Figure 1: Points used to label vole via DeepLabCut (DLC):** Each point was manually placed on both the male and female at the time of establishing ground truth. Different point colors were used to distinguish between the animals in each frame.

#### 4. Autoencoder

The results generated by the DLC network post training comprised of three varieties of errors: swapping, lagging, and shrinking. In order to keep the identities of the male and female separate, the network used two different skeletal colors: blue for female and red for the male. However, on certain occasions, especially when the two animals engaged in a series of rapid investigative or mating behaviors that resulted in their bodies

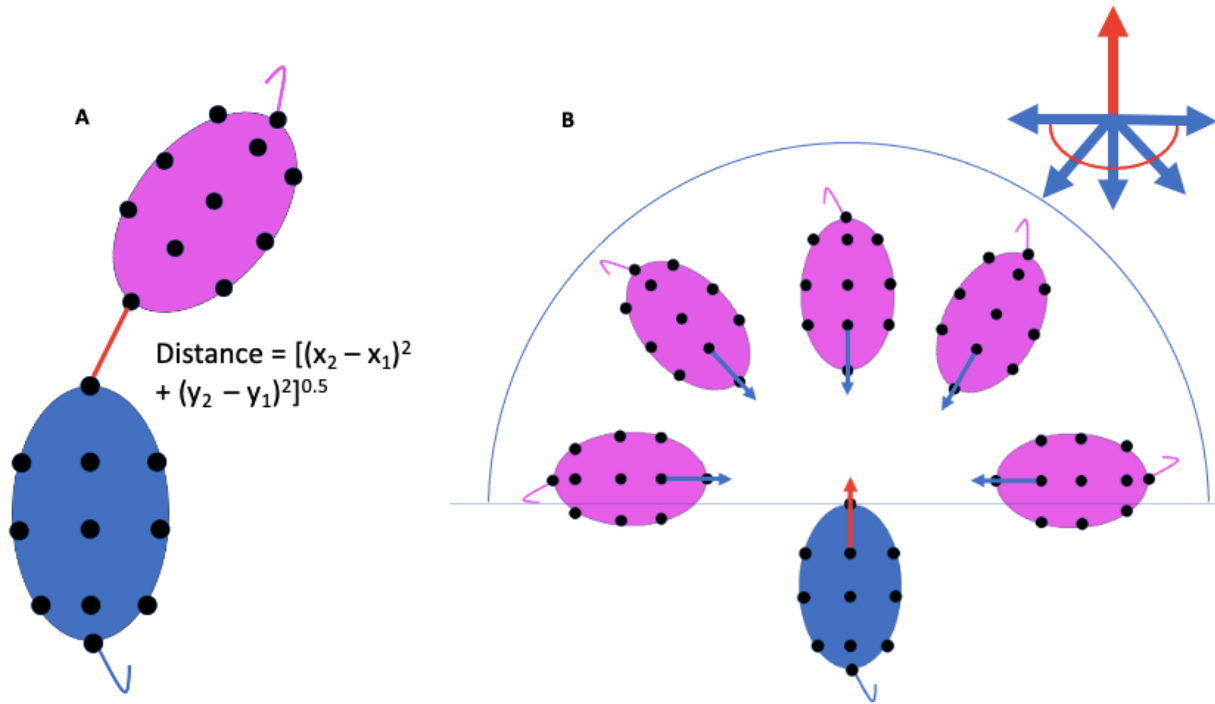
adopting complex configurations, the network failed to accurately maintain their identities. In other words, the skeletons “swapped” between the male and female. The “lagging” problem occurred on several occasions wherein an animal changed its position rapidly (for example, by accelerating very quickly from rest). As a result, the network failed to generate skeletons that were precisely placed on the animal’s body. On separate occasions, particularly when an animal adopted postures that made its shape difficult to decipher (i.e. accurately differentiate anterior from posterior or ventral from lateral), the “shrinking” problem persisted, wherein the skeletons reduced significantly in size and failed to map out the animal’s entire body. While each of these three types of errors was noticed across multiple videos, their occurrence is somewhat excusable given the fact that a human annotator would have also found these complex postures/movements difficult to accurately decipher. However, as such errors would have certainly produced faulty results and given rise to inaccurate conclusions regarding the differences between the WT and KO groups, it was extremely important to rectify them prior to analyzing the raw tracking data. We developed an autoencoder capable of identifying these badly labeled frames and repositioning skeletal points that deviated significantly from a predefined “shape” of the voles used in the study. Consisting of four primary layers (input, encoder, decoder, output), the autoencoder relied on a Gaussian mixture model (Fig. 11) to filter out any severe variations in the data, thereby significantly improving the DLC tracking results.

## 5. Data Analysis

Having obtained coordinates of the animals through the tracking process, the next step was to use the coordinate data to make various calculations pertaining to the relative separation between the animals and ultimately identify various social behaviors. As previously described, performing manual annotation via BORIS allows for the quick generation of annotation summaries: the total number of instances (or total duration) of a specific behavior included in the ethogram can easily be determined using the tools provided by the software. However, in the case of our automated tracking network, the raw data was not in the form of specific labels from a predefined ethogram, but merely as a spreadsheet containing several thousand coordinates for each animal. We used MATLAB to write code (see Appendix) that could extract this raw data and translate it into a summary of distinct behavioral categories. The main categories of social behavior analyzed in this study were “social investigation” (including *nose-to-nose sniffing*, *side sniffing*, and *anogenital sniffing*) and “mating”. As previously discussed, the goal was to be able to standardize the definitions of these behaviors in order to facilitate generalizability in the results across and within labs.

To determine instances of *nose-to-nose sniffing*, we relied on two parameters: distance and direction. We calculated the distance between the nose of the male and the nose of the female and also found the angle between the male’s and female’s eyeline vectors (Figure 2). The *eyeline vector* represented the vector connecting the middle of the neck and the nose for each animal. We then proposed that the act of *nose-to-nose sniffing* would occur on any occasion wherein the distance between the male and female noses was less than a certain threshold distance and, simultaneously, the angle between the two

*eyeline vectors* was between a certain threshold of angles. In trying to determine a reasonable threshold for distance, we correlated the auto-tracking results with the manual results for a range of different distance values and looked for the value that yielded the highest correlation (or  $R^2$ ) with the manual tracking results. We discuss the rationale and limitations behind this process in more detail later (see Discussion and Figure 18). Based on the peak  $R^2$  value, the threshold distance was chosen to be 25 pixels (or approximately the distance between the *nose* and *between ears* coordinates on an average prairie vole). The threshold range of angles was chosen to be +/- 90 degrees (or the area in front of the female nose where *nose-nose sniffing* was most likely to take place (Figure 2B). In the case of *anogenital sniffing*, we proposed that this event would take place on any occasion wherein the distance between the male's *nose* and the female's *tail start* was less than a certain threshold distance (25 pixels) and, simultaneously, the angle between the male *eyeline vector* and the female *tail vector* (the vector represented by the line joining the female *tail start* coordinate and the female *midbody* coordinate) was within a certain threshold (+/- 90 degrees) (Figure 3).



**Figure 2: Determining instances of nose-nose sniffing:** (A) The distance between the noses of the male (blue) and female (pink) were calculated as the Euclidean distance between the nose coordinates on each animal. (B) The direction of sniffing was determined by the angle between the two eyeline vectors. The possible range of angles is represented by the illustration on the top right (red arrow = male, blue arrows = female).

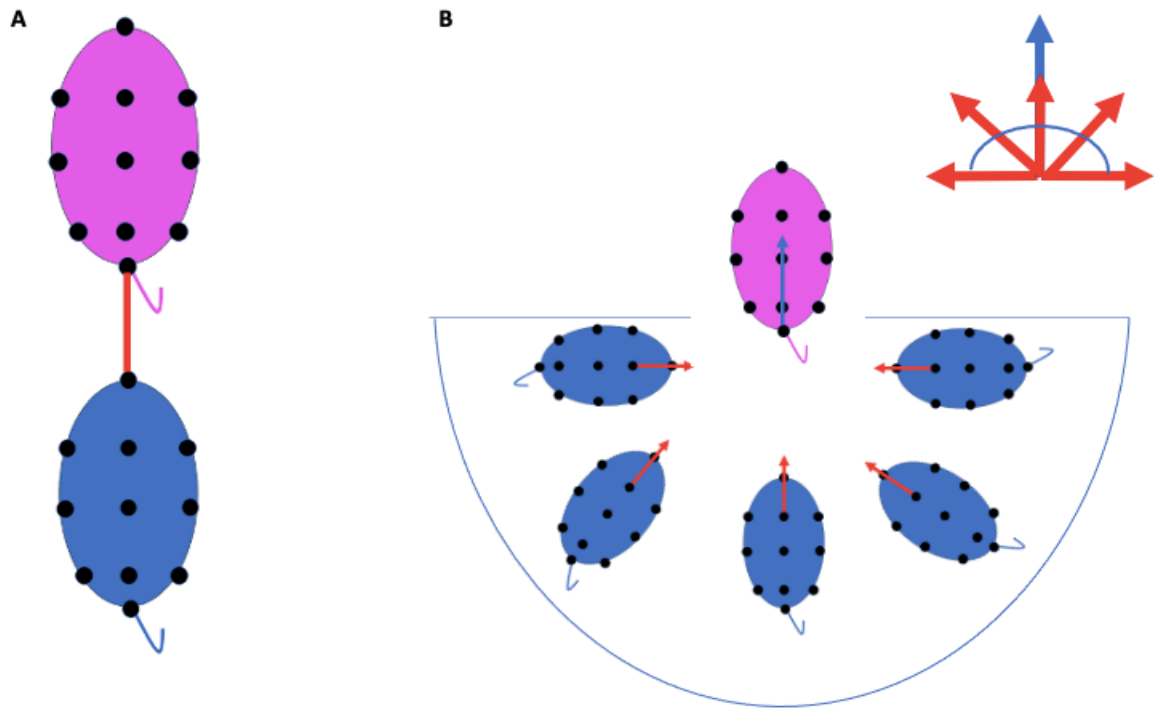
In determining instances of *side sniffing* (Figure 4), we sought to account for all cases wherein the male approached the female from either the right or left side of the female's body and went on to engage in the act of "sniffing". Firstly, we found the Euclidean distance between the male's nose and the female's midbody (d1), and between the male's nose and the female's (right/left) mid-waist (d2). Next, we also calculated the distance between the male's nose coordinate and the following segments on the female's body: ear to mid-waist (d3) and mid-waist to hip (d4) (for both the left and right sides of the body). We assumed that whenever d1 was greater than d2 (with reference to the *right-mid-waist*), then it was likely that the male was closer to the perimeter on the right

side of the female's body than to the center of the female's body (which would happen when the male *mounted* the female from the right side). Likewise, we assumed that whenever  $d_1$  was greater than  $d_2$  (with reference to the left-mid-waist), then it was likely that the male was closer to the perimeter of the left side of the female's body than to the center of the female's body (which would happen when the male *mounted* the female from the left side). In making these two assumptions, we were able to do two things: one, identify whether the male's nose was closer to the right or left side of the female body; and two, disregard instances wherein the male carried out a *sideward mounting* of the female (an act that would not constitute merely "sniffing" the female). Subsequently, we combined this knowledge pertaining to the male's position relative to the female's body sides with the aforementioned calculations concerning the distance between the male's nose coordinate and the body segments on the female. We stipulated that *side sniffing* would occur in the event that  $d_1$  was greater than  $d_2$  (for either the right or left sides of the female) and, simultaneously, the distance between the male's nose coordinate and either of the two segments ( $d_3/d_4$ ) (on the right or left sides of the female's body) was less than a predefined threshold distance.

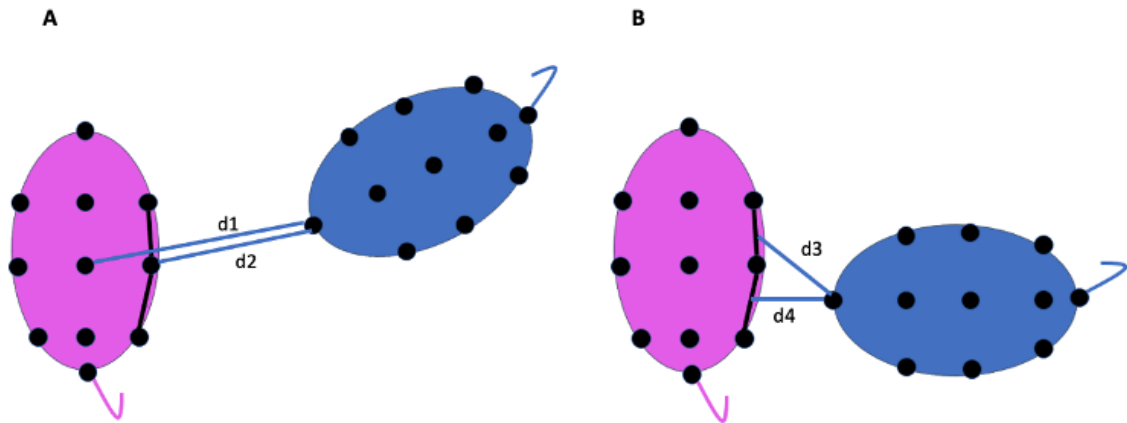
To identify instances of mating (Figure 5), it was important to come up with strict parameters for both distance as well as direction, since this behavior can only take place when the male and female are in a specific position relative to each other. We found two vectors: the first being the line joining the *midhip* and *midbody* coordinates on the male, and the second being the line joining the same two coordinates on the female. Next, similar to how we approached the calculation for *side sniffing*, we found the distance between the male *midbody* coordinate and the segment joining the female *midbody* and



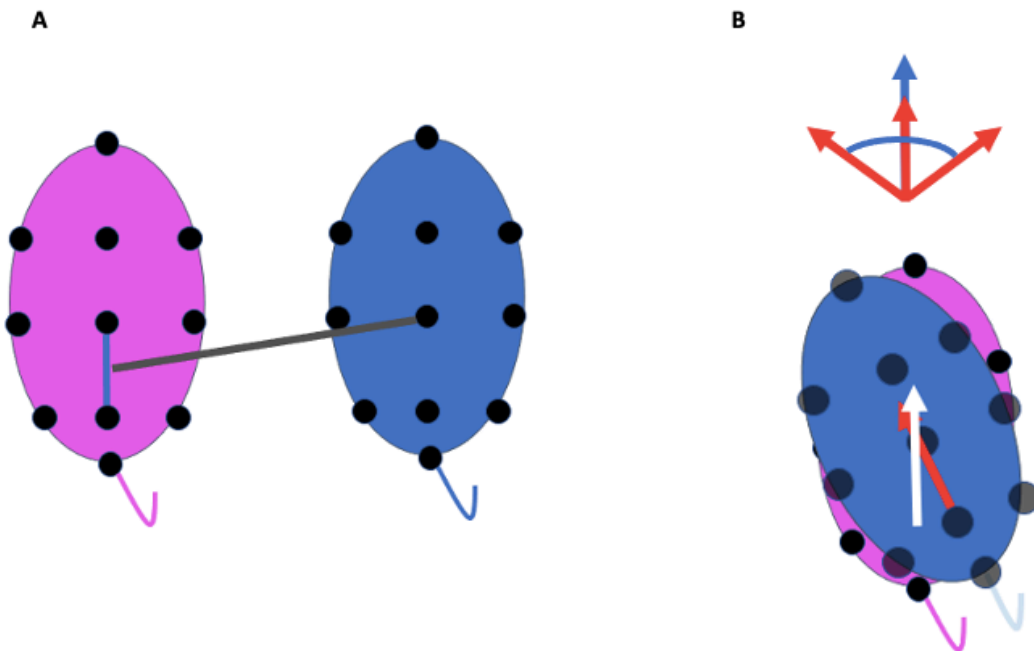
*midhip* coordinates. We assumed that in the event that this point-to-segment distance was less than a predefined threshold distance and, simultaneously, the angle between the two aforementioned vectors was less than a predefined threshold angle, the two animals would be engaged in *mating*.



**Figure 3: Determining instances of anogenital sniffing:** (A) The distance between the nose of the male (blue) and tail start of the female (pink) were calculated as the Euclidean distance between the nose and tail start coordinates on the male and female, respectively. (B) The direction of sniffing was determined by the angle between the male eyeline vector and the female tail vector. The possible range of angles is represented by the illustration on the top right (red arrows = male, blue arrow = female).



**Figure 4: Determining instances of side sniffing:** (A) The distance between the male nose coordinate and female midbody coordinate (d1) and the distance between the male nose coordinate and female (right/left) midwaist coordinate (d2) allowed for determination of whether the male was side-mounting or side-sniffing the female. (B) The distances between the male nose coordinate and female (right/left) *ear-midwaist* segment (d3) and the male nose coordinate and female (right/left) *midwaist-hip* segment (d4) were compared against the threshold distance value for *side sniffing*.



**Figure 5: Determining instances of mating:** (A) The distance between the male *midbody* coordinate and the segment joining the female *midbody* and *midhip* coordinates was compared against the predefined mating threshold. (B) The direction of mating was determined by

calculating the angle between the vectors joining the *midbody* and *midhip* coordinates on the bodies of the male and female. The white arrow (used for visibility) represents this vector on the female. The possible range of mating vectors is represented by the illustration on the top right. There, the white arrow has been replaced by the blue arrow.

## Results

### Comparing Auto-Tracked and Manually-Derived Results Independent of Genotype:

One of the key objectives of this study was to be able to create an automated tracking network that would not only reduce the time required for the annotation process, but also perform with equal, if not greater, accuracy than human annotators. Therefore, it was vital for us to compare the auto-tracking results side-by-side with the manual tracking results for specific behaviors across every animal. We compared manual and auto-tracked results for *mating*, and *social investigations* (*nose-to-nose sniffing*, *anogenital sniffing*, and *side sniffing*) across 5 WT and 6 Oxtr-KO cohabitations. It is worth noting here that in order to ensure greater validity of the auto-tracking results, we included in our code a conditional statement that disregarded instances wherein a behavior was observed for less than 0.5 seconds. Doing so was crucial because in the event that the network mistakenly identified, for example, multiple instances of *mating* (each lasting less than 0.5 seconds) while the animals were in fact engaged in another behavior like *anogenital sniffing*, these instances could ultimately add up to wrongly alter the final duration results for *mating*.

### Mating:

In comparing auto-tracked vs manually-generated results for the *total mating duration* during the first 30 minutes of cohabitation, we found that there was a strong positive correlation ( $R^2 = 0.86$ ,  $p = 0.42$  (t-test)) between the auto-tracked and manual annotation results (Figure 6). In other words, for each of the animals observed, the network that we developed was able to

capture a value for *total mating duration* that was almost equal to that captured by human annotators using BORIS. In addition to *duration*, we also calculated the *number of instances* captured by both the network and human annotators for each of the behaviors of interest. This was important to do because we wanted to see whether the durations being identified by the network were based on the same or different number of episodes of the behavior. In the case of *number of instances of mating* (Figure 8), we found a weak positive correlation ( $R^2 = 0.20$ ,  $p = 0.02$  (t-test)) between the auto- and manually-generated results. We offer a detailed explanation for this occurrence later (see Discussion) but we think that this difference has to do with the network's inability to differentiate between male- and female-initiated *mounting* behaviors, which tend to look nearly identical unless animal identities are clearly distinguished. The high positive slope value for the same results, however (Figure 8), indicates that, on average, the auto- and manually-generated calculations for *number of instances* did seem to line up when the network was able to correctly identify only male-initiated *mating* episodes.

### **Social Investigation:**

The human annotators did not differentiate between *nose-nose sniffing*, *anogenital sniffing*, and *side sniffing*, but simply grouped all three types of “sniffing” under the category of “social investigation.” The automated tracking network, on the other hand, was trained to identify each of these subtypes of *sniffing* as separate behaviors. In order to directly compare the results of the automated tracking network with those of manual annotation for *social investigation*, we simply found the sum of durations for *nose-nose sniffing*, *anogenital sniffing*, and *side sniffing* for each animal—as this value would be logically equal to the *social investigation* value determined by human annotators using BORIS. For all of the WT and KO animals studied, we found a positive correlation ( $R^2 = 0.63$ ,  $p = 0.74$  (t-test)) between the auto-

tracked and manually annotated results for the total duration of *social investigation* (Figure 7). Meanwhile, when comparing the *number of instances* of social investigation captured for each animal, we again found a positive correlation ( $R^2 = 0.62$ ,  $p = 0.85$  (t-test)) between the auto-tracked and manually annotated results (Figure 9).

### Comparing Quantitative Differences in Social Behavior Across Genotypes:

Having established the reliability of the network (i.e., its capability of generating results comparable to those generated by human annotators), we then sought to use it to examine certain quantitative differences in social behaviors displayed by WT and Oxt<sup>-/-</sup> male prairie voles when in the vicinity of their WT female conspecifics. We once again compared both *total duration* as well as *number of instances* for various social behaviors that are characteristic of prairie voles in cohabitation: *mating*, *nose-nose sniffing*, *anogenital sniffing*, and *side sniffing*. Additionally, for further validation of our auto-tracking results, we also compared these parameters between the two genotypic groups using manually-annotated data.

Using the results generated by our auto-tracking network, we found no significant differences in the *mean duration* (average total duration across all WT or all KO animals) between the WT and KO animals for each of the four behaviors investigated (Fig. 10). The following p values were calculated from the two-tailed t-test: 0.80 (*nose-nose sniffing*), 0.43 (*anogenital sniffing*), 0.72 (*side sniffing*), and 0.69 (*mating*). We also compared the mean *number of instances* for the same four behaviors across WT and KO animals and found no significant differences (Figure 13). The following p values were calculated from the two-tailed t-test: 0.93 (*nose-nose sniffing*), 0.42 (*anogenital sniffing*), 0.62 (*side sniffing*), and 0.20 (*mating*). For

reference, when we looked into these same comparisons using manually-generated data (Figure 12), we found similar results. There was no significant difference in the *mean duration* between the WT and KO animals for both *social investigation* and *mating*. The following p values were calculated from the two-tailed t-test: 0.94 (*social investigation*) and 0.63 (*mating*). Using manually-generated data, we also looked compared the mean *number of instances* for both *social investigation* and *mating* between the WT and KO animals and found no significant differences (Figure 15). The following p values were calculated from the two-tailed t-test: 0.91 (*social investigation*) and 0.95 (*mating*).

## Discussion

In this study, we trained a supervised machine learning network capable of automatically tracking prairie voles in cohabitation and thereby increasing the efficiency of the tracking process by nearly 100%. Moreover, our network sought to eliminate the variability and biases that often arise from manual tracking processes involving human annotators (Segalin et al., 2021). Our results showed that, when compared to manual tracking processes, the auto-tracking network was able to identify both the duration as well as frequency of various social behaviors with varying levels of precision, depending on the thresholds that we set. As indicated in Figures 6,7, and 9, at a threshold of 25 pixels (for both sniffing and mating behaviors), the  $R^2$  value is indicative of a strong correlation (about the mean) between the manually- and auto-derived tracking results. However, we found that upon exceeding a threshold of 25 pixels, or, likewise, decreasing the threshold below 25 pixels in descending order, the  $R^2$  values appear to decrease in both directions (Figure 18). This suggests that the number 25 serves as a peak threshold value (for sniffing behaviors) that allows for the generation of results that have minimal variance in the

context of our regression model. That said, upon setting the threshold value to 40 pixels, we can see that the slope of the correlation line (Figure 16) is exactly 1—indicating that the auto-tracked results for each animal are almost identical to the corresponding manually-generated results. However, we also observe that the y-intercept in this case (161.83) is larger than that seen for a lower threshold of 25 (101.91); and, in fact, the y-intercept value increases consistently with the slope value upon going from thresholds as low as 15 to as high as 40. Therefore, while it is tempting to aim for a threshold that provides a slope value of 1, the higher y-intercept values that accompany this higher slope value indicate longer durations of the behavior (social investigation) being detected by the auto-tracking network when compared to the manual tracking data. Upon using our code to look at the kinds of frames being extracted by the auto-tracking network at higher thresholds (40 pixels), we found that there were many “false positives” (Figure 19). By that we mean that the instances of *nose-nose sniffing*, for example, that the network detected, appeared to be too “distant” according to the threshold we would otherwise visually approximate during the process of manual annotation. This argument, however, assumes that our anthropocentric definition of what a “social interaction” should look like is accurate and applicable to prairie voles as well. The whole point of using the auto-tracking network was to be able to eliminate such biases in the behavioral annotation process and we will discuss this point further in the contents that follow. In addition to the above-mentioned occurrence of “false positives” at a threshold of 40 pixels, the  $R^2$  value at this threshold (0.55) also appears to be less than that for a threshold of 25 (0.63). The argument we pose in trying to choose between a higher slope value or a higher  $R^2$  value to show the robustness of our tracking network has to do with the significance and reliability of the “ground truth”. Of course, we took several measures to ensure that the manual annotation process was conducted as impartially as possible and

annotators were trained to carry out the labeling process with high accuracy. Even then, however, as discussed earlier, the manual annotation process will always be prone to errors (like different annotator response times and visual sensory abilities) that are beyond experimental control. As discussed earlier, when different experimenters or different labs have their own inherent biases, the question then boils down to whose ground truth is to be trusted and regarded as the “gold standard”. We propose that one of the advantages of our tracking network is that it offers experimenters the opportunity to choose their own threshold parameters, depending on their experimental conditions and the strictness with which they want to apply their behavioral definitions. Indeed, allowing for this flexibility ultimately takes away from the objectivity that this study aims to bring about in the field. That said, given the peak threshold values for sniffing behaviors that we were able to determine through our experiment, our model also offers experimenters the opportunity to identify optimal threshold values (based on peak  $R^2$  values or other parameters like slope or intercept offset) and to make this optimization explicit rather than implicit—as is the case with manual scoring. In doing so, those who implement our model might often find (as did we) that there exist significant differences in the manual and auto-tracked data. However, such differences might, in fact, be indicative of the biases or flaws present in the manual tracking methodology rather than of shortcomings in the accuracy of the auto-tracking network. In that sense, the earlier-mentioned argument regarding whose manually-labeled data is the gold standard might shift to, instead, whether auto-tracking itself is the real gold standard when compared to manual tracking. To provide statistical motivation for this question, it is worth considering Figure 17 at this point. When we look at the number of instances for all WT and KO animals wherein the animals are within different threshold distances of each other (the X-axis represents the pixel value for the distance between the male *nose* coordinate and the female



*midbody* coordinate), it appears that up until a threshold of 40 pixels, there are no significant differences between the two genotypes (which concurs with the findings of this study, given the thresholds that we implemented). However, upon going beyond a threshold of 40, all the way up until a threshold of 400 pixels, there seem to be many more instances wherein WT animals are in closer proximity than KO animals. This finding raises some interesting questions: does the communication involved in pair bond formation in prairie voles take place over longer distances than the those considered by the field thus far while performing manual scoring? If so, could the impairments resulting from the oxytocin receptor knockout actually manifest at higher distance thresholds than previously explored?

In Figure 8, we saw a moderately positive correlation ( $R^2 = 0.20$ ) between the auto-tracked and manually-annotated results for the *number of instances* of *mating*. In comparison to the correlation we found for the *number of instances* of *social investigation* ( $R^2 = 0.62$ ), this correlation appears to be much weaker. Given that the correlation between the auto-tracked and manually generated results for the *total duration* of the same behavior (*mating*) was quite strong ( $R^2 = 0.86$ ), meaning that both the network and human annotators observed similar *total mating durations* for the same animals, we propose that the weaker correlation observed for *number of instances* could be due to the network's inability to differentiate between instances wherein the male mounted the female from instances when the female mounted the male. As can be seen in Figure 6, there appear to be two cases wherein the auto-tracking results show large positive values (around 25 and 50) for *total mating duration* when the manual results are around 0 for the same two animals. After looking back at the specific frames that were identified by our auto-tracking network as instances of “mating”, we were able to confirm that a majority of these

frames involved the female mounting the male, an act that is not considered “mating”. In the case of manual annotation, such instances would quickly be filtered out given a human’s ability to distinguish between the male and female while watching the video. In the case of our auto-tracking model, however, given that the final behavioral annotations depend on the robustness or quality of behavioral definitions created by the human coder, there can often be instances wherein even good tracking results (by the network) fail to be highlighted appropriately (by human analyzers). While significant efforts were made to account for any positional nuances while coding the definitions for all other behaviors, including *nose-nose sniffing*, *anogenital sniffing*, and *side sniffing*, given the orientations of the male and female when engaging in *mating*, it was difficult for us to write the appropriate code to be able to differentiate between male- and female-initiated mounting episodes. However, given that mounting of the male by the female is relatively rare when compared to mounting of the female by the male (Blocker and Ophir, 2016), we do not see this as a matter of grave concern in the context of this study.

After having proven the reliability of our tracking network in generating results comparable to human annotators, we then sought to implement it in our exploration of the quantitative differences in the nature of social behaviors displayed by oxytocin receptor knockout male prairie voles when compared to wildtype male prairie voles. As indicated in Figures 10, 11, 12, 13, 14, and 15, we found no significant differences for either *total duration* or *number of instances* of various social behaviors between the WT and KO groups. This finding was common in both automated and manual annotation processes. In previously conducted studies involving partner preference tests, Oxt<sup>r</sup>-KO male prairie voles were shown to maintain their preferences towards their respective female partners (Wang and Aragona, 2004; Horie et

al., 2019). Based on those findings, we propose that it is likely that, in order to maintain such a preference, Oxt<sup>r</sup>-KO male prairie voles in cohabitation continue to display pro-social behaviors towards their female partners with similar frequency and duration as WT males. Furthermore, given that various other neurotransmitters and hormones, including vasopressin and dopamine, have previously been shown (Hammock and Young, 2006; Loth and Donaldson, 2021) to play a crucial role in the regulation of the pair bond (and hence the social behaviors governing pair bond formation), it is possible that the Oxt<sup>r</sup>-KO alone is insufficient to deter male prairie voles from engaging in pro-social behaviors towards their female partners with similar frequency and for similar durations as WT males.

On the whole, this study has demonstrated that it is possible to track multiple prairie voles in cohabitation and achieve results that are on par with those obtained from manual annotation processes. By analyzing the differences in social behavior enacted by Oxt<sup>r</sup>-KO and WT male prairie voles and finding the same outcomes through both automated and manual approaches, we are confident that our network can be further trained to eventually track multiple animals in three dimensions and thereby extract more nuanced behavioral patterns. In addition to displaying a high level of accuracy, our network reduces the time factor of the annotation process by nearly 100% and offers experimenters a unique level of flexibility with regards to adapting behavioral definitions and thresholds to suit the unique requirements of their studies.

## Future Directions

In this study we focused primarily on *mating* and *sniffing* behaviors to explore the quantitative differences in the nature of social interactions taking place across WT and Oxtr-KO male prairie voles. Another group of behaviors that we attempted to investigate, but were unable to successfully define for coding purposes in this project, is approach/departure behaviors. In addition to *sniffing* and *mating*, two very important characteristics of measuring affinity between conspecifics include the frequencies of approaches and departures—or simply how often the male approaches the female from a distance and how often the male moves away from the female after being in close proximity to it for an extended period of time, respectively. One of the challenges that we faced was in accounting for all possible mannerisms in which an approach or departure might take place. For instance, a male may run towards a female from rest at a high speed and in a straight line, or it may take a curved (or multi-direction) route to ultimately end up next to the female. In either case, however, the male may stop along the way at one or more occasions or simply adopt an uninterrupted motion. In such cases, it is unclear as to when exactly during the route (or journey towards the female) that the male receives the neurological intent to approach the female. Determining this intent can be challenging even for human annotators performing manual annotation, and hence there is a need for greater investigation into this topic. Additionally, as discussed earlier, upon looking at the cumulative distribution of frames across different threshold values for nose-body distance, we found that WT animals appear to be in closer proximity to each other than do KO animals at thresholds exceeding 40 pixels. This finding urged us to consider the possibility that the social interactions that oxytocin signaling facilitates during pair bond formation might in fact take place over much longer distances than previously expected. In future studies, we think it would be useful and important to amend our

definitions for social investigations and interactions to be able to account for these larger threshold values. The motivation to explore this phenomenon might not have arisen if not for the flexibility offered by our automated tracking model to select very specific threshold values and compare the subsequent outputs very quantitatively. Additionally, given the reliability of our network thus far, we hope to build on it further so that other researchers may implement it on other model organisms as well. Additionally, as discussed earlier, we hope to develop it to the extent where it can track animal movement in three dimensions. This will allow for a more nuanced analysis of behavior and potentially eliminate many of the tracking errors that result from the limitations of a two-dimensional camera angle.

## REFERENCES

- Bales, K. L. et al. What is a pair bond? *Horm Behav.* 136, 105062 (2021).
- Baribeau DA & Anagnostou E. Oxytocin and vasopressin: linking pituitary neuropeptides and Their receptors to social neurocircuits. *Front. Neurosci.* (9), e00335 (2015).
- Beery, A. K. Familiarity and Mate Preference Assessment with the Partner Preference Test. *Current Protocols* (1), e173 (2021).
- Berendzen, K. M. et al. Oxytocin receptor is not required for social attachment in prairie voles. *Neuron.* (111), 787-796 (2023).
- Blocker TD & Ophir AG. A preference to bond? Male prairie voles form pair bonds even in the presence of multiple receptive females. *Anim Behav.* (122), 89-97 (2016).
- Burkett J, Andari E, & Young L. Oxytocin-dependent consolation behavior in rodents. *Science* (351), 375-378 (2016).
- Fletcher, G. J. O. et al. Pair-Bonding, Romantic Love, and Evolution: The Curious Case of *Homo Sapiens*. *Perspectives on Psychological Science* 10 (1), 20–36 (2015).
- Hammock EAD & Young LJ. Oxytocin, vasopressin and pair bonding: implications for autism. *Philos Trans R Soc Lond B Biol Sci.* (361), 2187-2198 (2006).
- Heidenreich M & Zhang F. Applications of CRISPR-Cas systems in neuroscience. *Nat Rev Neurosci.* (17), 36-44 (2016).
- Horie, K. et al. Oxytocin receptor knockout prairie voles generated by CRISPR/Cas9 editing Show reduced preferences for social novelty and exaggerated repetitive behaviors. *Horm Behav.* (111), 60-69 (2019).
- Horie K, Inoue K, Nishimori, K, & Young L. Investigation of Oxtr-expressing Neurons

- Projecting to Nucleus Accumbens using Oxytr-ires-Cre Knockout in prairie Voles (*Microtus ochrogaster*). *Neuroscience* (448), 312-324 (2020).
- Johnson ZV & Young L. Neurobiological mechanisms of social attachment and pair bonding. *Curr Opin Behav Sci.* 3, 38-44 (2015).
- Josipović, G. et al. Antagonistic and synergistic epigenetic modulation using orthologous CRISPR/dCas9-based modular system. *Nucleic Acids Research* (47), 9637-9657 (2019).
- Kitano, K. et al. Helping behavior in prairie voles: A model of empathy and the importance of oxytocin. *iScience* (25), 103991 (2022).
- Liu Y & Wang Z. Nucleus accumbens oxytocin and dopamine interact to regulate pair bond formation in female prairie voles. *Neuroscience* (121), 537-544 (2003).
- López-Gutiérrez, M. F. et al. Brain functional networks associated with social bonding in monogamous voles. *eLife* (10), e55081 (2021).
- Mathis, A. et al. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience* (21), 1281-1289 (2018).
- Meredith KL & Donaldson ZR. Oxytocin, Dopamine, and Opioid Interactions Underlying Pair Bonding: Highlighting a Potential Role for Microglia. *Endocrinology* (162), bqaa223 (2021).
- Nilsson, S. RO. et al. Simple Behavioral Analysis (SimBA) – an open source toolkit for computer classification of complex social behaviors in experimental animals.
- Noldus LP, Spink AJ, & Tegelenbosch RA. EthoVision: a versatile video tracking system for Automation of behavioral experiments. *Behav Res Methods Instrum Comput.* (33), 398-414 (2001).
- Pereira, T. D. et al. SLEAP: Multi-animal pose tracking. *Nature Methods* (19), 486-495 (2022).

- Rae, M et al. Oxytocin and vasopressin: Signaling, behavioral modulation and potential Therapeutic effects. *British Journal of Pharmacology* (179), 1544-1564 (2021).
- Ross, H. E. et al. Variation in Oxytocin Receptor Density in the Nucleus Accumbens has Differential Effects on Affiliative Behaviors in Monogamous and Polygamous Voles. *Journal of Neuroscience* (29), 1312-1318 (2009).
- Sánchez-Macouzet, O. et al. Better stay together: pair bond duration increases individual fitness Independent of age-related variation. *Proc Biol Sci.* 281(1786) 20132843 (2014).
- Segalin, C. et al. The Mouse Action Recognition System (MARS) software pipeline for automated analysis of social behaviors in mice. *eLife* (10), e63720 (2021).
- Shemesh, Y. et al. High-order social interactions in groups of mice. *eLife* (2), e00759 (2013).
- Silverman JL, Yang M, Lord C, & Crawley JN. Behavioral phenotyping assays for mouse models of autism. *Nat Rev Neurosci.* (11), 490-502 (2010).
- Song Z & Albers HE. Cross-talk among oxytocin and arginine-vasopressin receptors: Relevance For basic and clinical studies of the brain and periphery. *Front Neuroendocrinol.* (51), 14-24 (2017).
- Sturman, O. et al. Deep learning based behavioral analysis enables high precision rodent tracking and is capable of outperforming commercial solutions. *Neuropsychopharmacology* (45), 1942-1952 (2020).
- Walter T & Couzin ID. TRex, a fast multi-animal tracking system with markerless identification, and 2D estimation of posture and visual fields. *eLife* (10), e64000 (2021).
- Wang Z & Aragona BJ. Neurochemical regulation of pair bonding in male prairie voles. *Physiology & Behavior* (83), 319-328 (2004).
- Williams JR, Carter CS, & Insel T. Partner preference development in female prairie voles is



Facilitated by mating or the central infusion of oxytocin. *Ann N Y Acad Sci.* (652), 487-489 (1992).

Wiltschko, A. B. et al. Mapping Sub-Second Structure in Mouse Behavior.

*Neuron* (88), 1121-1135 (2015).

*BioRxiv* (2020).

Winslow, J. T. Mouse social recognition and preference. *Curr Protoc Neurosci.* (8), 8.16 (2003).

Young L & Wang Z. The neurobiology of pair bonding. *Nat Neurosci* (7), 1048-1054 (2004).

## APPENDIX

### MATLAB Codes

#### Script A:

```

close all; clear all; clc;
%% This script is going to read .xlsx files storing automatic tracking
results
folderAddress='/Users/rahilriadmahmood/Documents/MATLAB/Track_cohab'; %%
where the xls files storing the tracking results are saved
DataList=dir([folderAddress,'/Corrected_Tracking_Results']);
for ff=1:size(DataList,1)-2
    filename=DataList(ff+2).name;
    strings=split(filename,'_');
    for tt=1:size(strings,1)
        if strfind(strings{tt}, 'male')
            recordingName=strings{tt};
        else
            end
    end
    if strfind(filename, 'H0')
        NameTag='H0_';
    else
        NameTag='WT_';
    end
    [num, text, raw] = xlsread([folderAddress,'/Corrected_Tracking_Results/',
filename]);

    coord_data(1).nose = num(:, 2:3);
    coord_data(2).nose = num(:, 24:25);
    coord_data(1).leftear = num(:, 4:5);
    coord_data(2).leftear = num(:, 26:27);
    coord_data(1).betweenears = num(:, 6:7);
    coord_data(2).betweenears = num(:, 28:29);
    coord_data(1).rightear = num(:, 8:9);
    coord_data(2).rightear = num(:, 30:31);
    coord_data(1).rightmidwaist = num(:, 10:11);
    coord_data(2).rightmidwaist = num(:, 32:33);
    coord_data(1).midbody = num(:, 12:13);
    coord_data(2).midbody = num(:, 34:35);
    coord_data(1).leftmidwaist = num(:, 14:15);
    coord_data(2).leftmidwaist = num(:, 36:37);
    coord_data(1).lefthip = num(:, 16:17);
    coord_data(2).lefthip = num(:, 38:39);
    coord_data(1).midhip = num(:, 18:19);
    coord_data(2).midhip = num(:, 40:41);
    coord_data(1).righthip = num(:, 20:21);
    coord_data(2).righthip = num(:, 42:43);
    coord_data(1).tailstart = num(:, 22:23);
    coord_data(2).tailstart = num(:, 44:45);
    targetpath=[folderAddress,'/Mat_files_of_tracking_results/'];
    if ~exist(targetpath)
        mkdir(targetpath);

```

```

    else
    end
    save([targetpath, NameTag, recordingName, '.mat'], 'coord_data');
end

```

### Script B:

```

close all; clear all; clc;
%% load the .mat files saving tracking results one by one in the data folder,
% decide at each frames whether the male animals (animal 1) are involved in
% 1. nose-to-nose sniffing; 2. anogenital sniffing; 3. side sniffing;
% 4. mating; 5. approaching; 6. departing (definition of 5 and 6
unfinished)
% and the results are saved into matrix of index and saved in
% BehavIndexData.mat
%% Direct to the root folder of the project containing children data folders
folderAddress='/Users/rahilriadmahmood/Documents/MATLAB/Track_cohab/'; %%
where the xls files storing the tracking results are saved
%% Settings of threshold
threshold_nn_si = 15; % threshold for nose-nose sniffing
threshold_ag_si = 15; % threshold for ano-genital sniffing
threshold_ns_si = 15; % threshold for side sniffing
threshold_mating = 25; % threshold distance for mating

mating_angle = 45;
nn_angle_min = +90; % nose-nose sniffing angle
ag_angle_max = 90; % ano-genital sniffing angle

DataList=dir([folderAddress, 'Mat_files_of_tracking_results']);
for ff=1:size(DataList,1)-3 %% !!!!!!!!!!!!!
    filename=DataList(ff+3).name; %%!!!!!!!!!!!!!!
    load([folderAddress, '/Mat_files_of_tracking_results/', filename]);
    BehavIndex(ff).recFile=filename;
    %%
    loc_n = size((coord_data(1).nose),1);
    Index = zeros(loc_n, 6); % store indexes indicating behavioral status for
each behavioral ethogram
    % 1: nose-2-nose sniffing; 2: anogenital sniffing; 3: side sniffing;
    % 4: mating; 5: approaching; 6: departing

    %%
    n2bodyDis=zeros(loc_n,2);
    for ii = 1:loc_n
        M_nose = coord_data(1).nose(ii,:); % extracting all male nose
coordinates from raw data
        M_midbody = coord_data(1).midbody(ii,:); % extracting all male
midbody coordinates from raw data
        M_midhip = coord_data(1).midhip(ii,:); % extracting all male midhip
coordinates from raw data
        M_betweenears = coord_data(1).betweenears(ii,:); % extracting all
male between ears coordinates from raw data
        F_betweenears = coord_data(2).betweenears(ii,:); % extracting all
female between ears coordinates from raw data
    end
end

```

```

        F_nose = coord_data(2).nose(ii,:); % extracting all female nose
coordinates from raw data
        F_tail = coord_data(2).tailstart(ii,:); % extracting all female tail
start coordinates from raw data
        F_EarR = coord_data(2).rightear(ii,:); % extracting all female right
ear coordinates from raw data
        F_MidWaistR = coord_data(2).rightmidwaist(ii,:); % extracting all
female right mid waist coordinates from raw data
        F_HipR = coord_data(2).righthip(ii,:); % extracting all female right
hip coordinates from raw data
        F_EarL = coord_data(2).leftear(ii,:); % extracting all female left
ear coordinates from raw data
        F_MidWaistL = coord_data(2).leftmidwaist(ii,:); % extracting all
female left mid waist coordinates from raw data
        F_HipL = coord_data(2).lefthip(ii,:); % extracting all female left
hip coordinates from raw data
        F_BetEars = coord_data(2).betweenears(ii,:); % extracting all female
betweenears coordinates from raw data
        F_MidHip = coord_data(2).midhip(ii,:); % extracting all female midhip
coordinates from raw data
        F_MidBody = coord_data(2).midbody(ii,:); % extracting all female
midhip coordinates from raw data

        %% finding instances of nose-nose sniffing
        n2nDis(ii,1) = pdist([M_nose; F_nose], 'euclidean'); % finding
distance between male nose and female nose
        MHeadVec = M_nose - M_betweenears; % defining calculation for male
head vector
        FHeadVec = F_nose - F_betweenears; % defining calculation for female
head vector
        anglehn2hn(ii,1) = VecAngle(MHeadVec , FHeadVec); % calculating angle
between male and female head vectors for nose-nose investigation
        if n2nDis(ii,1) < threshold_nn_si & abs(anglehn2hn(ii,1)) >
nn_angle_min
            Index(ii,1) = 1;
        else
            end

        %% finding instances of ano-genital sniffing
        n2tsDis(ii,1) = pdist([M_nose; F_tail], 'euclidean'); % finding
distance between male nose and female tail start
        FTailVec = F_tail - F_MidHip; % defining calculation for female tail
vector
        anglehn2tn(ii,1) = VecAngle(MHeadVec, FTailVec); % calculating angle
between male head vector and female tail vector for ano-genital investigation
        if n2tsDis(ii,1) < threshold_ag_si & abs(anglehn2tn(ii,1)) <
ag_angle_max
            Index(ii,2) = 1;
        else
            end

        %% finding mating instances
        MMatingVec = M_midbody - M_midhip; % defining calculation for male
mating vector

```

```

    FMatingVec = F_MidBody - F_MidHip; % defiing calculation for female
    mating vector
    anglemh2mb(ii,1) = VecAngle(MMatingVec, FMatingVec); % calculating
    angle between male and female mating vectors for mating_instances
    [dis2, status2]=PointToSegDist(M_midbody, F_MidBody, F_MidHip);
    if dis2 < threshold_mating & abs(anglemh2mb(ii,1)) < mating_angle
        Index(ii,4) = 1;
    else
    end

    %% side sniffing parameters calculation
    if pdist([M_nose; F_MidBody]) > pdist([M_nose; F_MidWaistL]) %
    male's nose is closer to the left side
        [dis, status]=nose2bodyDis(M_nose, F_EarL, F_MidWaistL, F_HipL);
        n2bodyDis(ii,1)=dis;
        n2bodyDis(ii,2)=status;
        if n2bodyDis(ii,1)<threshold_ns_si
            if status == 2
                Index(ii,3)=1;
            else
                if sum(Index(ii,[1,2,4]))
                else
                    Index(ii,3)=1;
                end
            end
        end
    else
    end
    elseif pdist([M_nose; F_MidBody]) > pdist([M_nose; F_MidWaistR]) %
    male's nose is closer to the right side
        [dis, status]=nose2bodyDis(M_nose, F_EarR, F_MidWaistR, F_HipR);
        n2bodyDis(ii,1)=dis;
        n2bodyDis(ii,2)=status;
        if n2bodyDis(ii,1)<threshold_ns_si
            if status == 2
                Index(ii,3)=1;
            else
                if sum(Index(ii,[1,2,4]))
                else
                    Index(ii,3)=1;
                end
            end
        end
    else
    end
end

end

%% calculating approach/departure instances (not finished!!!)
fs=20; %% frame per second for the videos
frleng= 11; % for Sgolay filtering
order = 3; % for Sgolay filtering

male_mb = coord_data(1).midbody; %% coordinates of male's midbody
female_mb = coord_data(2).midbody; %% coordinates of female's midbody

```

```

velocity_male=diff(male_mb,1)*fs;
Vm2f=female_mb-male_mb;
% % velocity filtered for smoothing with the Savitzky-Golay filtering
% smoothV=sgolayfilt(velocity_male, order, frleng);
% for jj=1:size(smoothV,1)
%     smoothSpeed(jj,1) = norm(smoothV(jj,:));
%     smoothHeading_angle(jj, :) = VecAngle(smoothV(jj,:), Vm2f(jj, :));
% end
%
% acceleration_male=diff(smoothSpeed);
% norm_acceleration=zscore(acceleration_male);
%%
% [FragNum Frags]=ContinAcc(norm_acceleration, 3, 2);
% %%
% figure
% time1=(1:1:jj)/20;
% bx(1)=subplot(4,1,1)
% plot(time1,smoothSpeed, 'k');
% xlabel('Time (s)');
% ylabel('Velocity (pixel/s)');
% title('Velocity');
%
% bx(2)=subplot(4,1,2)
% plot(time1(1:end-1), norm_acceleration); hold on;
% for ff=1:FragNum
%     text(Frags(ff)/fs, norm_acceleration(ff),'o','color','r')
% end
% title('Normalized accelerator');
%
% bx(3)=subplot(4,1,3)
% plot(time1,smoothSpeed, 'k');
% findpeaks(smoothSpeed, time1, 'MinPeakProminence', 100);
% [pks locs w p] = findpeaks(smoothSpeed, time1, 'MinPeakProminence',
100);
% title('Find peaks in velocity data');
%
% bx(4)=subplot(4,1,4)
% plot(time1, smoothHeading_angle, 'r');
% xlabel('Time');
% ylabel('Direction of female');
% title('Angle between movement and position vectors');
% linkaxes(bx,'x');
% xlim([0 3600]);
% size(pks)
% %%
% figure
% polarhistogram(smoothHeading_angle/pi, 16*4);
% title('Distribution of angles between female and movement directions');
% %%
% figure
% scatter(smoothHeading_angle(1:end-1), norm_acceleration, 'r.');
```

hold on

```

plot([-180 180], [3 3], 'b--'); hold on
plot([-180 180], [-3 -3], 'b--');
xlim([-180 180]);
xlabel('Angles between female and movement directions');
```

```

%     ylabel('Normalized accelerator');

%%
%
% idxSN = find(side_sniffing(:,3));
% idxNN = find(n2tsDis(:,2));
% idxAG = find(anogen_sniffing_instances(:,1));
% idxMating = find(mating_instances(:, 1));
% plotN=10*6;
% downN=ceil(length(idxAG)/plotN);
% % plotFrameVectors(coord_data, idxAG(1:downN:end))
BehavIndex(ff).Index=Index;
clearvars coord_data n2nDis anglehn2hn n2tsDis anglehn2tn anglemh2mb
n2bodyDis
end
%%
save([folderAddress, 'BehavIndexData.mat'], "BehavIndex")

%% Calculate angle between two given vectors, result is in in degree
function angle = VecAngle(vector1, vector2)
angle = atan2d(vector1(1)*vector2(2)-vector1(2)*vector2(1),
vector1(1)*vector2(1)+vector1(2)*vector2(2));
end

%% Plot the skeletons of animals with coord_data and frame index
function plotFrameVectors(data, frameIdx)
number = length(frameIdx);
if number <= 6
    figure
    for plotNum = 1:number
        subplot(2,3,plotNum)
        ff = frameIdx(plotNum);
        SubPlotVectors(data, ff)

        xlim([0 640]);
        ylim([0 480]);
    end
else
    for figNum = 1:ceil(number/6)    %%% 6 subplots per figure
        figure
        for plotNum = 1:min(6, number-6*(figNum-1))
            subplot(2,3,plotNum)
            idx = 6*(figNum-1)+plotNum;
            ff = frameIdx(idx);
            SubPlotVectors(data, ff)
            xlim([0 640]);
            ylim([0 480]);
        end
    end
end
end
end

function SubPlotVectors(data, ii)

```

```

x01 = [data(1).nose(ii,1), data(1).leftear(ii,1), data(1).leftmidwaist(ii,1),
data(1).lefthip(ii,1), data(1).tailstart(ii,1),...
data(1).righthip(ii,1), data(1).rightmidwaist(ii,1),
data(1).rightear(ii,1), data(1).nose(ii,1),...
data(1).betweenears(ii,1), data(1).midbody(ii,1), data(1).midhip(ii,1)];
y01 = [data(1).nose(ii,2), data(1).leftear(ii,2), data(1).leftmidwaist(ii,2),
data(1).lefthip(ii,2), data(1).tailstart(ii,2),...
data(1).righthip(ii,2), data(1).rightmidwaist(ii,2),
data(1).rightear(ii,2), data(1).nose(ii,2),...
data(1).betweenears(ii,2), data(1).midbody(ii,2), data(1).midhip(ii,2)];
xI1 = [data(1).leftear(ii,1), data(1).betweenears(ii,1),
data(1).rightear(ii,1)];
yI1 = [data(1).leftear(ii,2), data(1).betweenears(ii,2),
data(1).rightear(ii,2)];
xII1 = [data(1).leftmidwaist(ii,1), data(1).midbody(ii,1),
data(1).rightmidwaist(ii,1)];
yII1 = [data(1).leftmidwaist(ii,2), data(1).midbody(ii,2),
data(1).rightmidwaist(ii,2)];
xIII1 = [data(1).lefthip(ii,1), data(1).midhip(ii,1),
data(1).righthip(ii,1)];
yIII1 = [data(1).lefthip(ii,2), data(1).midhip(ii,2),
data(1).righthip(ii,2)];

x02 = [data(2).nose(ii,1), data(2).leftear(ii,1), data(2).leftmidwaist(ii,1),
data(2).lefthip(ii,1), data(2).tailstart(ii,1),...
data(2).righthip(ii,1), data(2).rightmidwaist(ii,1),
data(2).rightear(ii,1), data(2).nose(ii,1),...
data(2).betweenears(ii,1), data(2).midbody(ii,1), data(2).midhip(ii,1)];
y02 = [data(2).nose(ii,2), data(2).leftear(ii,2), data(2).leftmidwaist(ii,2),
data(2).lefthip(ii,2), data(2).tailstart(ii,2),...
data(2).righthip(ii,2), data(2).rightmidwaist(ii,2),
data(2).rightear(ii,2), data(2).nose(ii,2),...
data(2).betweenears(ii,2), data(2).midbody(ii,2), data(2).midhip(ii,2)];
xI2 = [data(2).leftear(ii,1), data(2).betweenears(ii,1),
data(2).rightear(ii,1)];
yI2 = [data(2).leftear(ii,2), data(2).betweenears(ii,2),
data(2).rightear(ii,2)];
xII2 = [data(2).leftmidwaist(ii,1), data(2).midbody(ii,1),
data(2).rightmidwaist(ii,1)];
yII2 = [data(2).leftmidwaist(ii,2), data(2).midbody(ii,2),
data(2).rightmidwaist(ii,2)];
xIII2 = [data(2).lefthip(ii,1), data(2).midhip(ii,1),
data(2).righthip(ii,1)];
yIII2 = [data(2).lefthip(ii,2), data(2).midhip(ii,2),
data(2).righthip(ii,2)];
plot(x01,y01, 'b-'); hold on % plot male's skeleton
plot(xI1,yI1, 'b-'); hold on % plot male's skeleton
plot(xII1,yII1, 'b-'); hold on % plot male's skeleton
plot(xIII1,yIII1, 'b-'); hold on % plot male's skeleton
plot(x02,y02, 'm-'); hold on % plot female's skeleton
plot(xI2,yI2, 'm-'); hold on % plot female's skeleton
plot(xII2,yII2, 'm-'); hold on % plot female's skeleton
plot(xIII2,yIII2, 'm-'); hold on % plot male's skeleton
title(ii);
end

```



```

%% calculating the shortest distance from a point to a segment
function [dis, statusIdx] = PointToSegDist(pt, p1, p2)
V1=pt-p1;
V2=pt-p2;
V0=p2-p1;
angle1 = VecAngle(V1, V0);
angle2 = VecAngle(V2, -V0);
if abs(angle1) > 90
    dis = pdist([pt; p1]);
    statusIdx = 1;
elseif abs(angle2) > 90
    dis = pdist([pt; p2]);
    statusIdx = 3;
else
    dis=sin(abs(angle1)/180*pi)*pdist([pt;p1]);
    statusIdx = 2;
end
end

%% distance between male's nose and body segments of the female
function [dis, statusIdx2] = nose2bodyDis(nose, ear, midwaist, hip)
[dis1,con1] = PointToSegDist(nose, ear, midwaist);
[dis2,con2] = PointToSegDist(nose, midwaist, hip);
if con1 == 1
    statusIdx2 = 1;
    dis = dis1;
elseif con2 == 3
    statusIdx2 = 3;
    dis = dis2;
else
    statusIdx2 = 2;
    dis = min(dis1,dis2);
end
end

```

### Script C:

```

close all; clear all; clc;
%%
%% Direct to the root folder of the project containing children data folders
folderAddress='./'; %% where the xls files storing the tracking results are
saved
load([folderAddress, 'BehavIndexData.mat']);
annoFiles=dir([folderAddress, 'Annotation_results']);
%% Parameters to be modulated
fps = 20;
minthreshold = 0.5; %% recognized behavioral bouts last less than
minthreshold will be excluded
ProcT = 30; %% Process behaviors *min after the female introduced in
%% Use two arrays to store the index of H0 and WT animals
idxH0=[];

```

```

idxWT=[];
for ii=1:size(BehavIndex,2)
    if strfind(BehavIndex(ii).recFile, 'H0')
        idxH0=[idxH0; ii];
    elseif strfind(BehavIndex(ii).recFile, 'WT')
        idxWT=[idxWT; ii];
    else
        end
    end
end
%% Store animal names in the file names of .xls files storing annotations
for jj=1:size(annoFiles,1)-2
    strings=split(annoFiles(jj+2).name, '_');
    AnnoDataName(jj).animal=strings{3};
end
%% Calculate bout time of each behaviors in an animal-by-animal way
behav={'n2n_sniffing', 'anogenital_sniffing', 'side_sniffing', 'mating',
'approach', 'departure'};
for ii=1:size(BehavIndex,2)
    StatByAnimal(ii).recFile=BehavIndex(ii).recFile;
    Index=BehavIndex(ii).Index;
    % Read the right excel file storing annotations and find the start time
    % when female is introduced in
    for jj=1:size(AnnoDataName, 2)
        if strfind(BehavIndex(ii).recFile, AnnoDataName(jj).animal)
            StatByAnimal(ii).annoIdx=jj;
            [num, text, raw] = xlsread([folderAddress, 'Annotation_results/',
annoFiles(jj+2).name]);
            annoT=num(2:end,10:12);
            anno=text(2:end,7);
            startT=annoT(strcmp(anno, 'female in'),1);
            %% write to calculate mating behaviors in 30min
            AnnoMatT1=annoT(find(strcmp(anno, 'mount only')),:);
            [AnnoMatDur1, AnnoMatNum1] = BehavWinCount(AnnoMatT1, [startT
startT+ProcT*60]);
            AnnoMatT2=annoT(find(strcmp(anno, 'mouting')),:);
            [AnnoMatDur2, AnnoMatNum2] = BehavWinCount(AnnoMatT2, [startT
startT+ProcT*60]);
            %% SI Calculation
            AnnoSIT=annoT(find(strcmp(anno, 'social investigation')),:); %%
            ###
            [AnnoSIDur, AnnoSINum] = BehavWinCount(AnnoSIT, [startT
startT+ProcT*60]); %% ###
            else
                end
            end
            StatByAnimal(ii).StartT=startT;
            StatByAnimal(ii).AnnoMatDur_window=AnnoMatDur1+AnnoMatDur2;
            StatByAnimal(ii).AnnoMatNum_window=AnnoMatNum1+AnnoMatNum2;
            StatByAnimal(ii).AnnoSIDur_window=AnnoSIDur-(AnnoMatDur1+AnnoMatDur2);
            %% ###
            StatByAnimal(ii).AnnoSINum_window=AnnoSINum;

            for kk=1:size(Index,2)
                behavT=ContiBehav(Index(:,kk), fps, minthreshold, startT);
                StatByAnimal(ii).(behav{1,kk})=behavT;
            end
        end
    end
end

```

```

        clearvars behavT
    end
    AllSniff=[StatByAnimal(ii).(behav{1,1}); StatByAnimal(ii).(behav{1,2});
StatByAnimal(ii).(behav{1,3})]; %% ###
    StatByAnimal(ii).TotalSniffing_Dur=sum(AllSniff(:,3)); %% ###
    StatByAnimal(ii).TotalSniffing_Num=size(AllSniff,1); %% ###
    clearvars Index num text raw annoT anno AllSniff %% ###
end

save([folderAddress, 'StatByAnimal.mat'], 'StatByAnimal');

%%
for ii=1:size(BehavIndex(1).Index, 2)
    StatByBehaviors(ii).behav=behav{1,ii};
    StatByBehaviors(ii).HO_duration_window = [];
    StatByBehaviors(ii).HO_number_window = [];
    StatByBehaviors(ii).WT_duration_window = [];
    StatByBehaviors(ii).WT_number_window = [];
    for jj=1:size(StatByAnimal,2)
        behavT=StatByAnimal(jj).(behav{1,ii});
        [TrackBehavDur, TrackBehavNum] = BehavWinCount(behavT, [startT
startT+Proct*60]);
        if strfind(StatByAnimal(jj).recFile, 'H0')
            StatByBehaviors(ii).HO_duration_window =
[StatByBehaviors(ii).HO_duration_window; TrackBehavDur];
            StatByBehaviors(ii).HO_number_window =
[StatByBehaviors(ii).HO_number_window; TrackBehavNum];
        elseif strfind(StatByAnimal(jj).recFile, 'WT')
            StatByBehaviors(ii).WT_duration_window =
[StatByBehaviors(ii).WT_duration_window; TrackBehavDur];
            StatByBehaviors(ii).WT_number_window =
[StatByBehaviors(ii).WT_number_window; TrackBehavNum];
        else
            end
        clearvars behavT
    end
end
StatByBehaviors(ii+1).behav = 'Total sniffing'; %% ###
StatByBehaviors(ii+1).HO_duration_window =
sum([StatByBehaviors(1).HO_duration_window, StatByBehaviors(2).HO_duration_win
dow, StatByBehaviors(3).HO_duration_window],2); %% ###
StatByBehaviors(ii+1).HO_number_window =
sum([StatByBehaviors(1).HO_number_window, StatByBehaviors(2).HO_number_window,
StatByBehaviors(3).HO_number_window],2); %% ###
StatByBehaviors(ii+1).WT_duration_window =
sum([StatByBehaviors(1).WT_duration_window, StatByBehaviors(2).WT_duration_win
dow, StatByBehaviors(3).WT_duration_window],2); %% ###
StatByBehaviors(ii+1).WT_number_window =
sum([StatByBehaviors(1).WT_number_window, StatByBehaviors(2).WT_number_window,
StatByBehaviors(3).WT_number_window],2); %% ###

save([folderAddress, 'StatByBehaviors.mat'], 'StatByBehaviors');
%% Summarize mating duration and numbers from annotation and tracking
AnnoMatDur=[];
AnnoMatNum=[];

```

```

for ii=1:size(StatByAnimal,2)
AnnoMatDur=[AnnoMatDur;StatByAnimal(ii).AnnoMatDur_window];
AnnoMatNum=[AnnoMatNum;StatByAnimal(ii).AnnoMatNum_window];
end
TrackMatDur=[StatByBehaviors(4).H0_duration_window;StatByBehaviors(4).WT_dura
tion_window];
TrackMatNum=[StatByBehaviors(4).H0_number_window;StatByBehaviors(4).WT_number
_window];
figure
subplot(1,2,1)
scatter(AnnoMatDur, TrackMatDur, 'r');
xlabel('Duration from Manual Annotation');
ylabel('Duration from Auto-Tracking');
title('Mating Duration in First 30 mins of Cohab')
R_dur=corrcoef(AnnoMatDur, TrackMatDur); %% ###
text(40, 160, ['r = ',num2str(R_dur(1,2))]); %% ###
subplot(1,2,2)
scatter(AnnoMatNum, TrackMatNum, 'r');
xlabel('Number of Events from Manual Annotation');
ylabel('Number of Events from Auto-Tracking');
title('Number of Mating Events in First 30 mins of Cohab')
R_num=corrcoef(AnnoMatNum, TrackMatNum); %% ###
text(5, 30, ['r = ',num2str(R_num(1,2))]); %% ###
%%
function behavT=ContiBehav(IndexArray, fps, minthreshold, startT)
temptP=IndexArray;
idxSig=find(temptP);
idxSig(find(idxSig<(startT*fps)))=[]; %% exclude all events before female
introduced in
FragNum=0;
FragS=[];
if isempty(idxSig)
else
    %%%% if any significant data point
    dis=diff(idxSig);
    idxSigJump=find(dis>1);
    if isempty(idxSigJump)
        %% if no jump of non-zero data points
        if diff([idxSig(1),idxSig(end)])<minthreshold*fps-1
            temptP(idxSig(1):idxSig(end))=0;
        else
            FragNum=FragNum+1;
            FragS=[idxSig(1) idxSig(end)];
        end
    else
        %%%% if there're jumps of non-zero data
points
        %% deal with the first Fragment of significant data points prior
        %% to the first jump
        if diff([idxSig(1),idxSig(idxSigJump(1))])<minthreshold*fps-1
            temptP(idxSig(1):idxSig(idxSigJump(1)))=0;
        elseif idxSig(1)==idxSig(idxSigJump(1))
            temptP(idxSig(1))=0;
        else
            FragNum=FragNum+1;
            FragS=[FragS,[idxSig(1) idxSig(idxSigJump(1)-1)]];
        end
    end
end

```

```

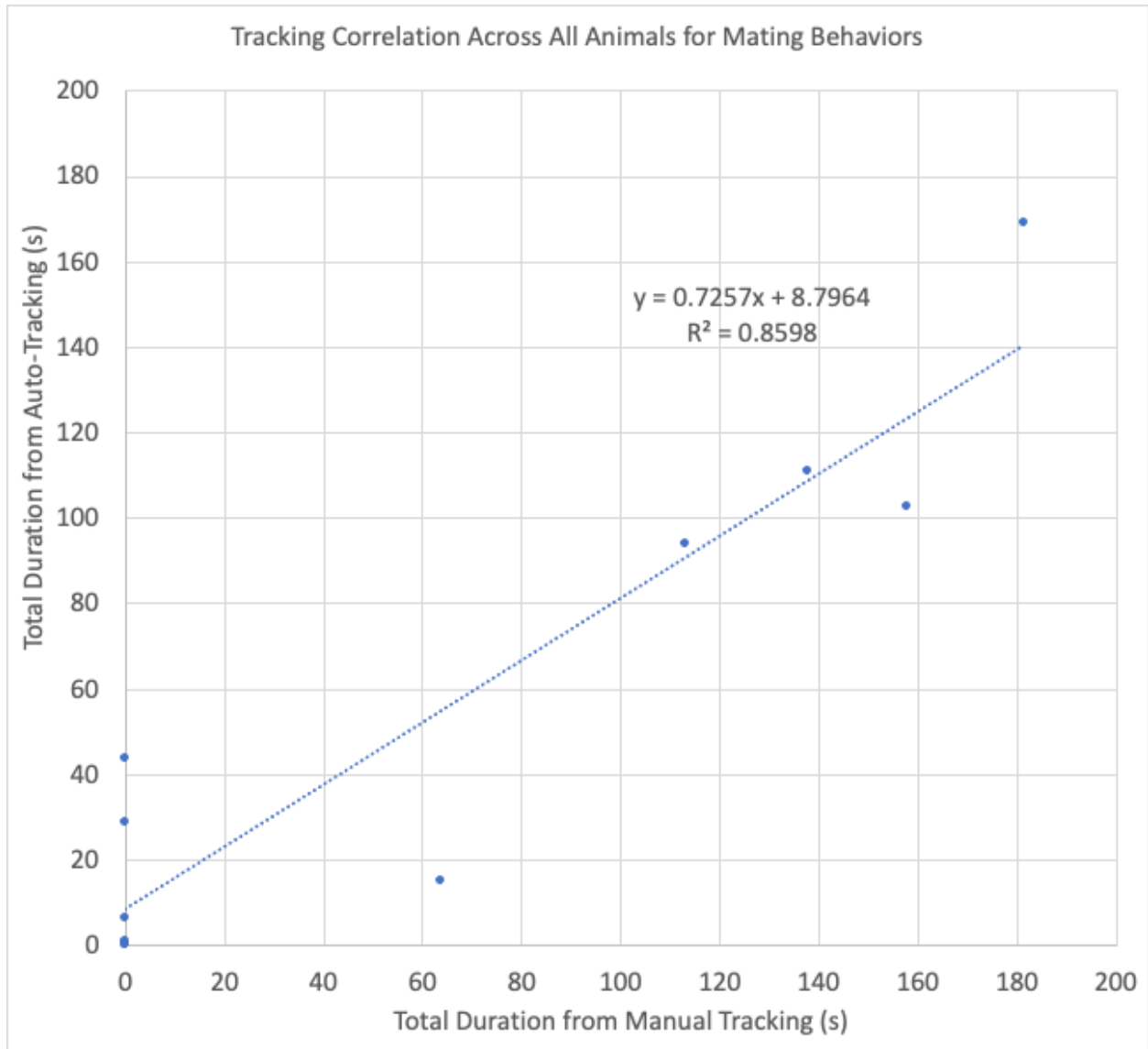
    %% deal with the other Fragments of significant data points
    for iii=1:length(idxSigJump) %%
        if iii<length(idxSigJump)
            if
diff([idxSig(idxSigJump(iii)+1),idxSig(idxSigJump(iii+1))])<minthreshold*fps-
1
            temptP(idxSig(idxSigJump(iii)+1):idxSig(idxSigJump(iii)))=0;
            else
                FragNum=FragNum+1;
                Frags=[Frags;[idxSig(idxSigJump(iii)+1)
idxSig(idxSigJump(iii+1))]];
            end
            else %% data after the last jump
                if diff([idxSig(idxSigJump(iii)+1),
idxSig(end)])<minthreshold*fps-1
                    temptP(idxSig(idxSigJump(iii)+1):idxSig(end))=0;
                else
                    FragNum=FragNum+1;
                    Frags=[Frags;[idxSig(idxSigJump(iii)+1) idxSig(end)]];
                end
            end
        end
    end
end
if FragNum==0
    Frags=[];
    behavT=[];
else
    behavT(:,[1 2])=Frags/fps;
    behavT(:,3)=behavT(:,2)-behavT(:,1);
    % behavT(find(behavT(:,3)<minthreshold),:)=[];
end
end
%%
function [duration, number] = BehavWinCount(behavT, window)
if ~isempty(behavT)
    idx=find(behavT(:,1)<window(2) & behavT(:,1)>window(1)); %% star time of
bouts located within the given time window
    if ~isempty(idx)
        number=length(idx);
        if behavT(idx(end),2)<window(2) %% if the end time of the last bout
located before time window too
            duration=sum(behavT(idx,3));
        elseif length(idx)>1 %% if there're more than one bout and the last
bout ends after window's end
            duration=sum(behavT(idx(1:end-
1),3))+diff([behavT(idx(end),1),window(2)]);
        else %% if there's only one bout and it ends after window's end
            duration=diff([behavT(idx(1),1),window(2)]);
        end
        if idx(1)>1
            if behavT(idx(1)-1,2)>window(1) %% consider if there's a bout
start before window's start but end after it

```

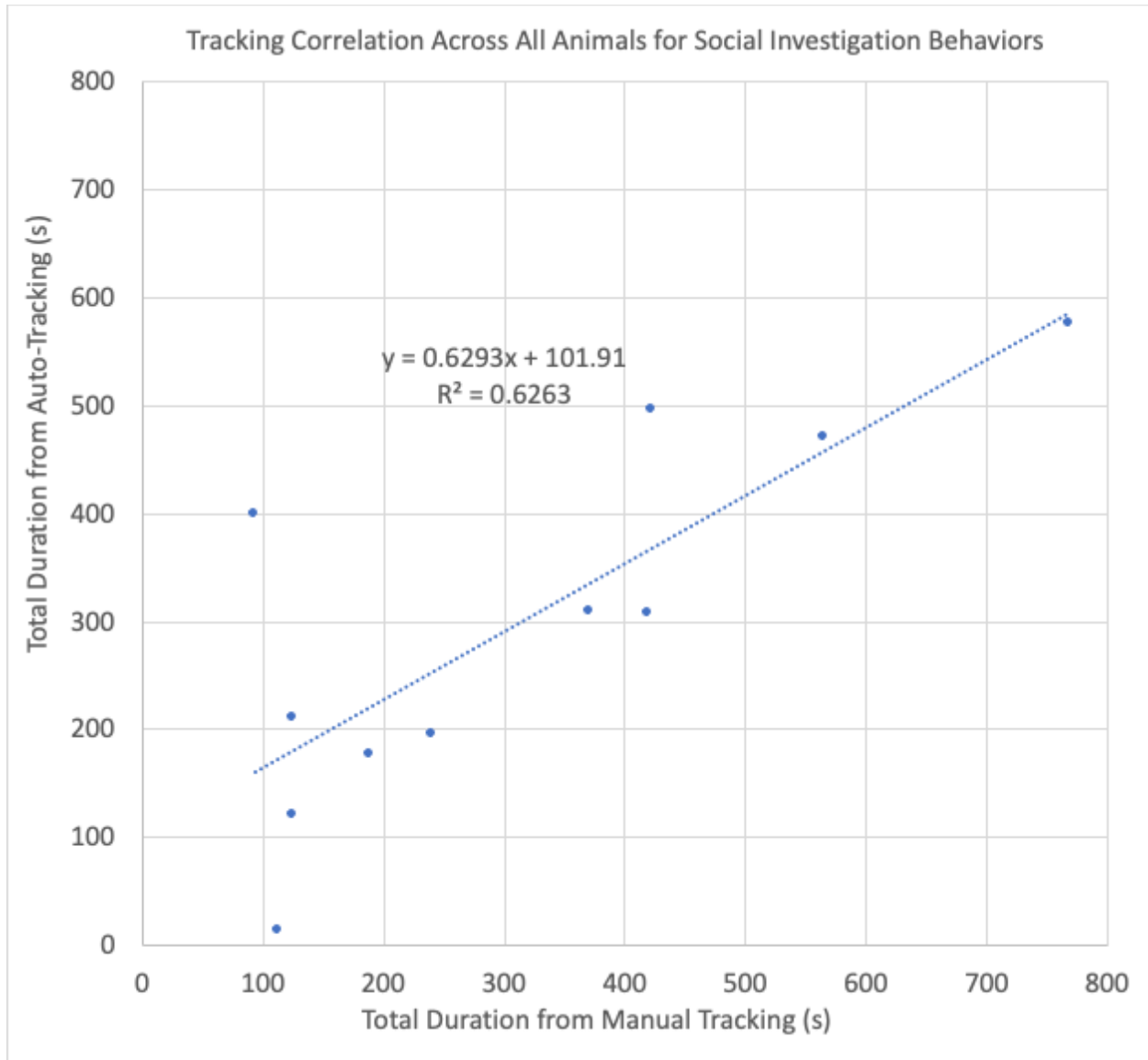
```
        duration=duration+diff([window(1), behavT(idx(1)-1,2)]);
    else
    end
else
end

else
    duration=0;
    number=0;
end
else %% no behavT input
    duration=0;
    number=0;
end
end
```

## FIGURES

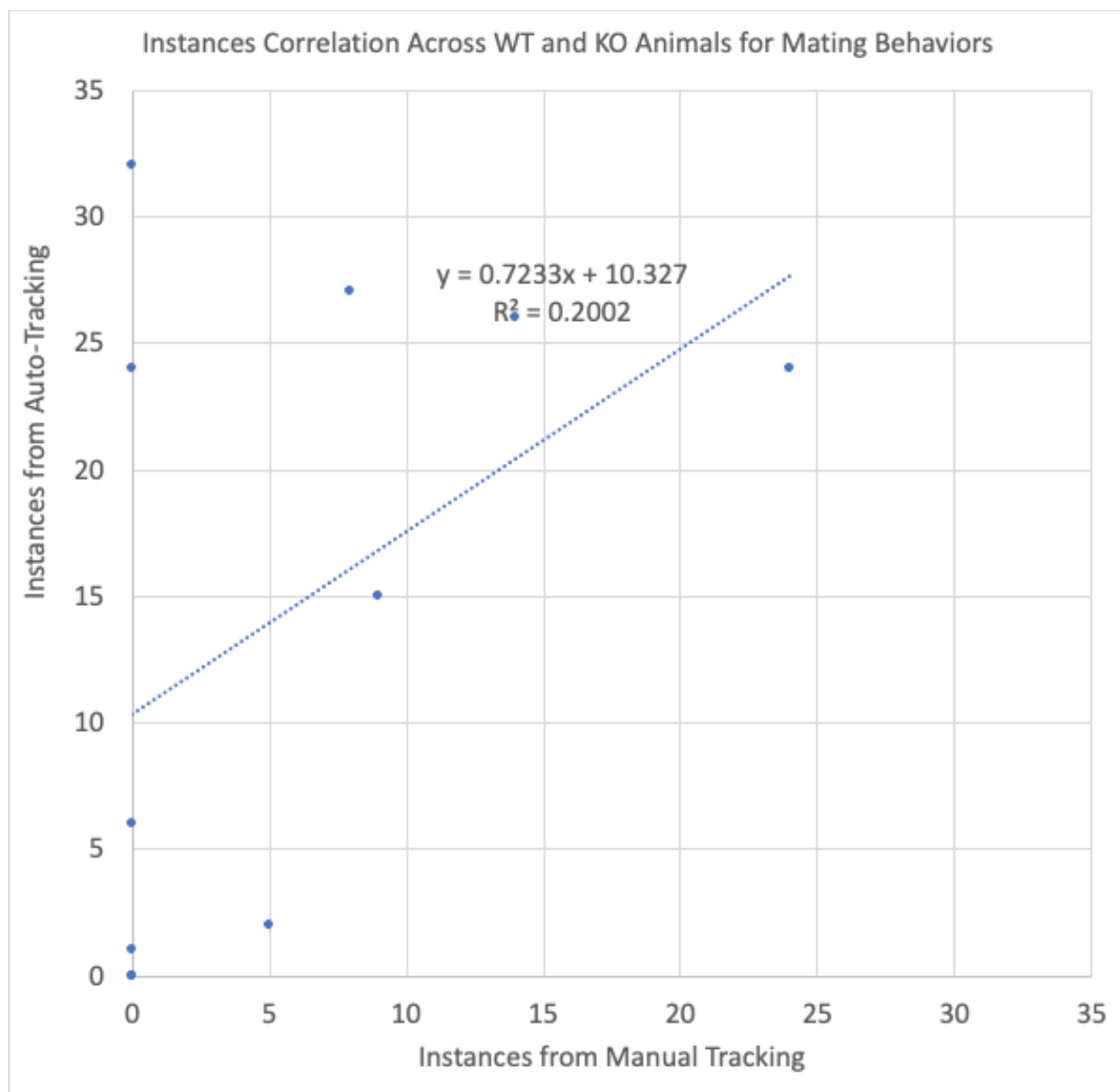


**Figure 6: Correlation between auto-tracked and manually generated results (mating duration).** The high  $R^2$  value indicates minimal difference in the total duration of mating captured by manual and automated tracking approaches. Each point represents a single animal ( $n=12$ ).

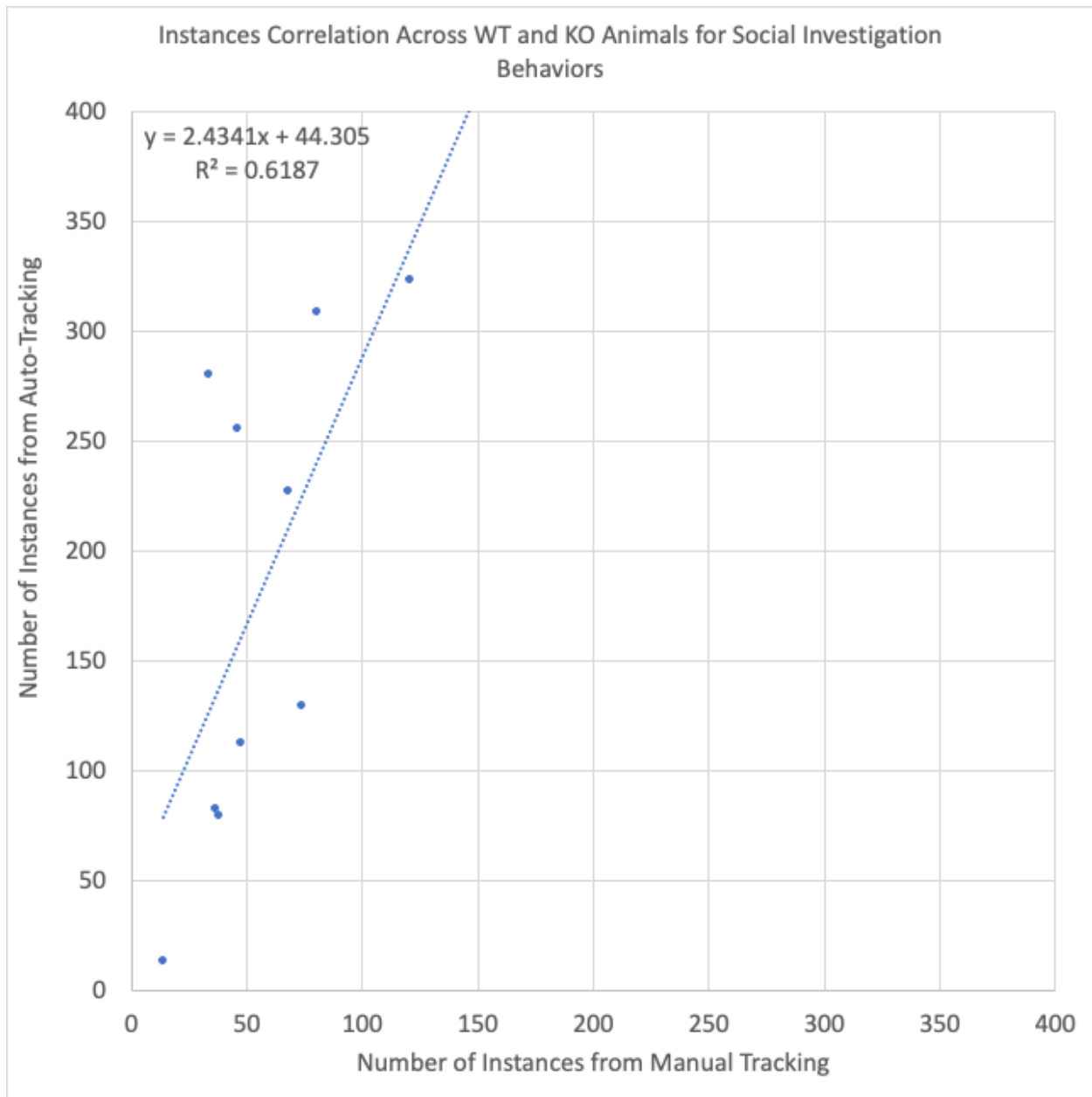


**Figure 7: Correlation between auto-tracked and manually generated results (social investigation duration).** The high  $R^2$  value indicates minimal difference in the total duration of social investigation captured by manual and automated tracking approaches. Each point represents a single animal ( $n=11$ ).

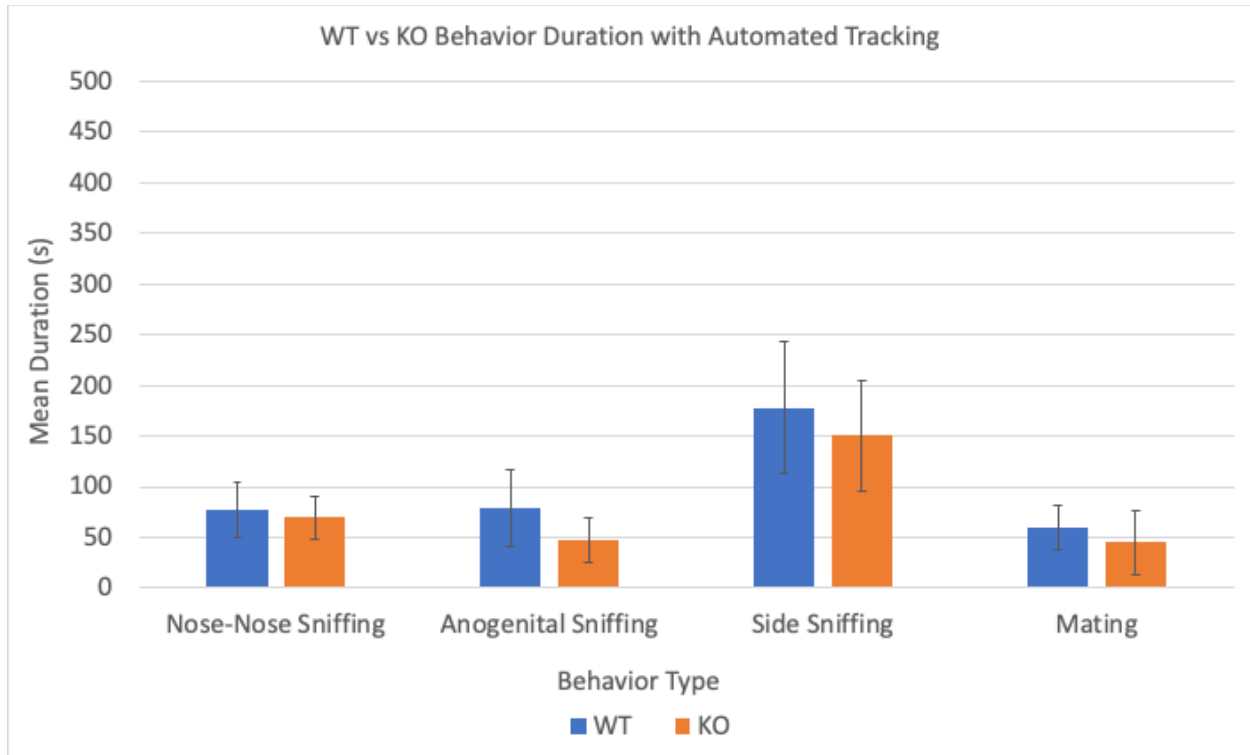




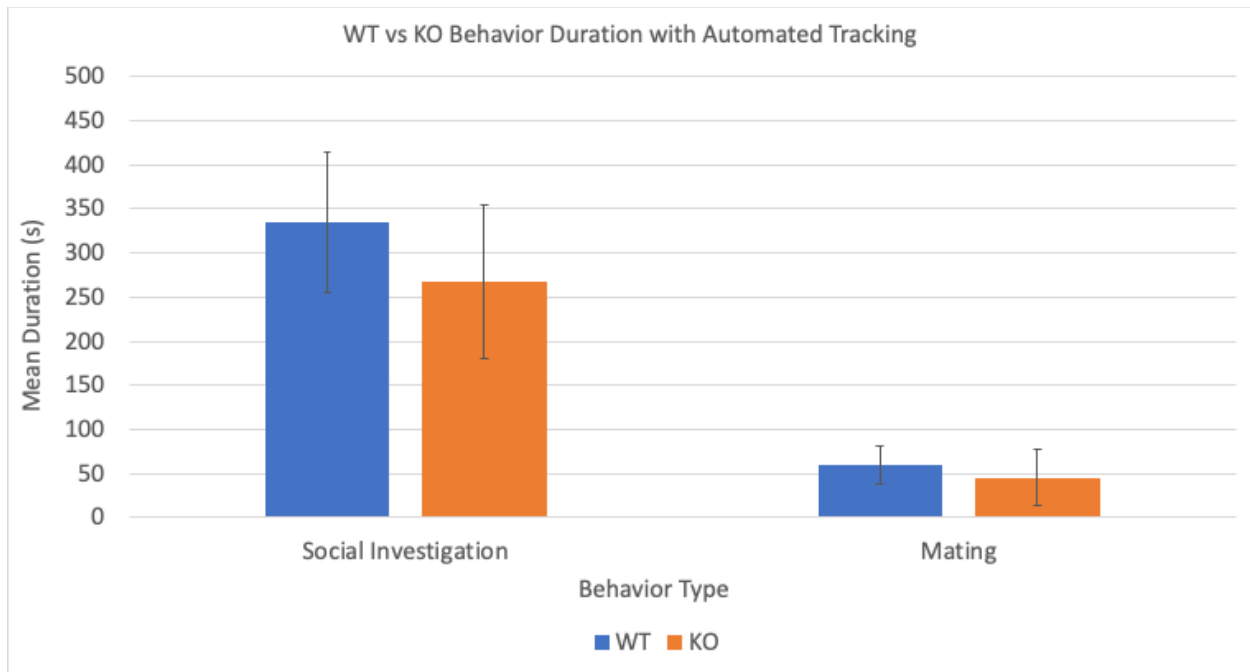
**Figure 8: Correlation between auto-tracked and manually generated results (mating instances).** The  $R^2$  value indicates moderate difference in the total number of instances of mating captured by manual and automated tracking approaches. Each point represents a single animal ( $n=12$ ).



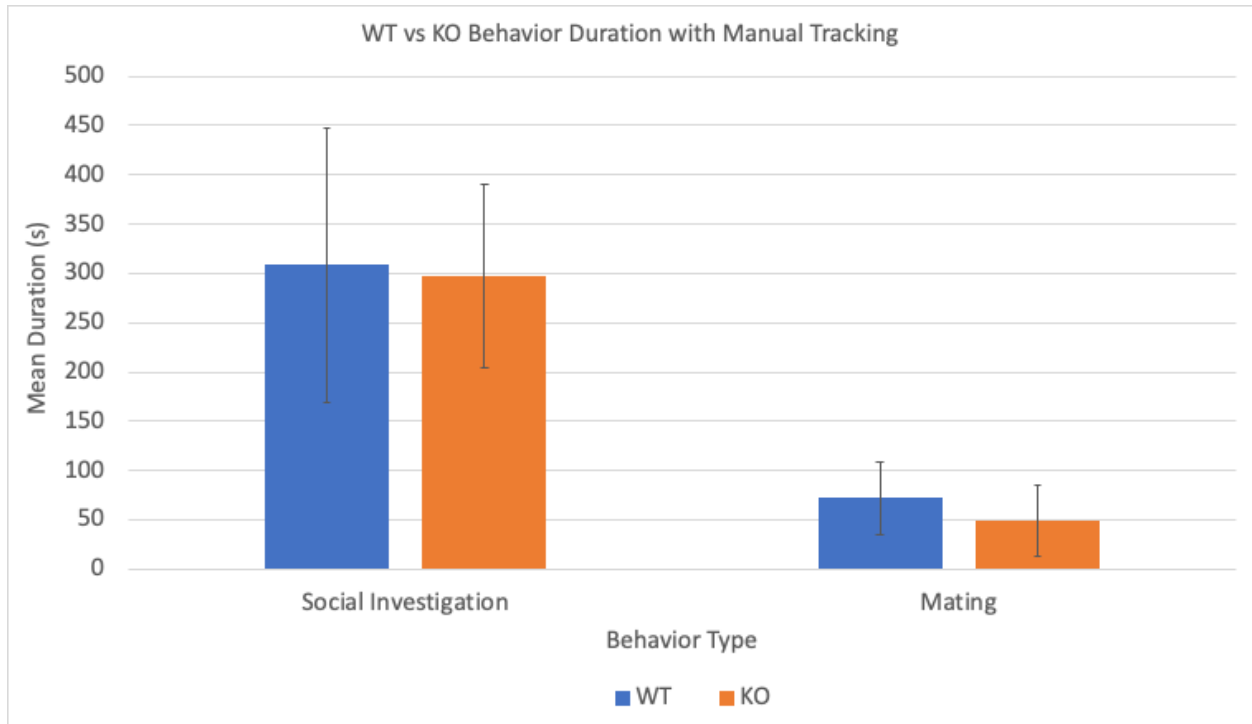
**Figure 9: Correlation between auto-tracked and manually generated results (social investigation instances).** The high  $R^2$  value indicates minimal difference in the total number of instances of social investigation captured by manual and automated tracking approaches. Each point represents a single animal ( $n=11$ ).



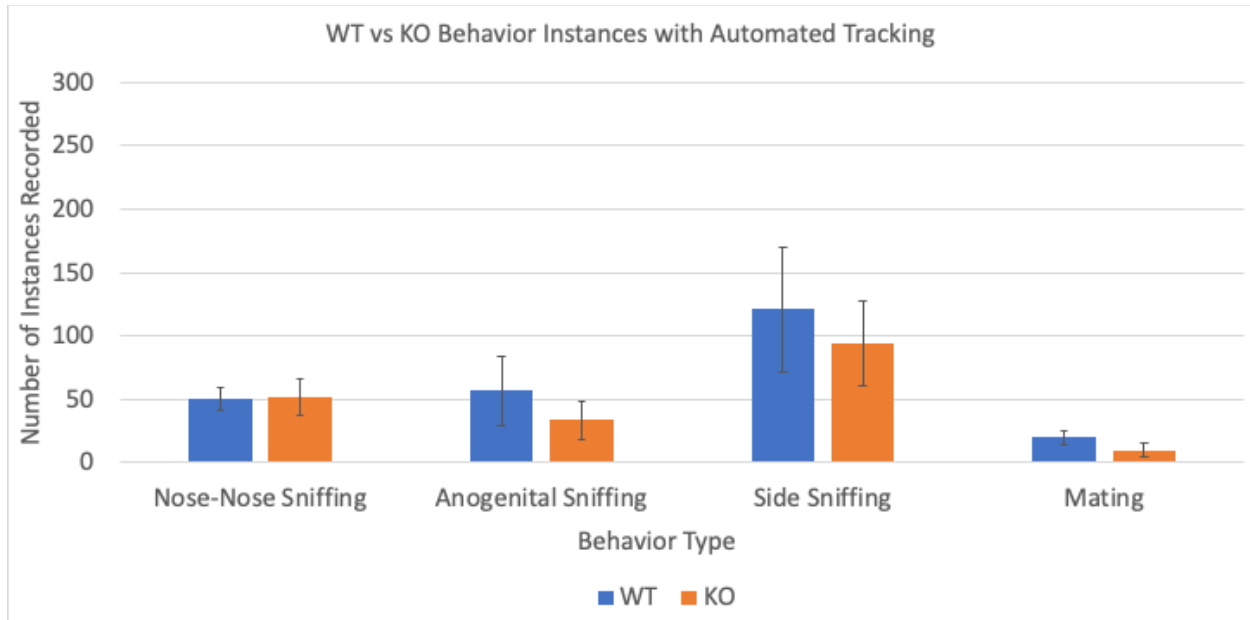
**Figure 10: Distribution of average durations of social behaviors (auto-tracking).** Each bar represents the average total time that the behavior was observed across all animals in the genotypic group (n=6 KO, n=5 WT). The differences between WT and KO for all four behaviors are insignificant (t-test, 2-tailed, unpaired: 0.73, 0.58, 0.89, 0.59).



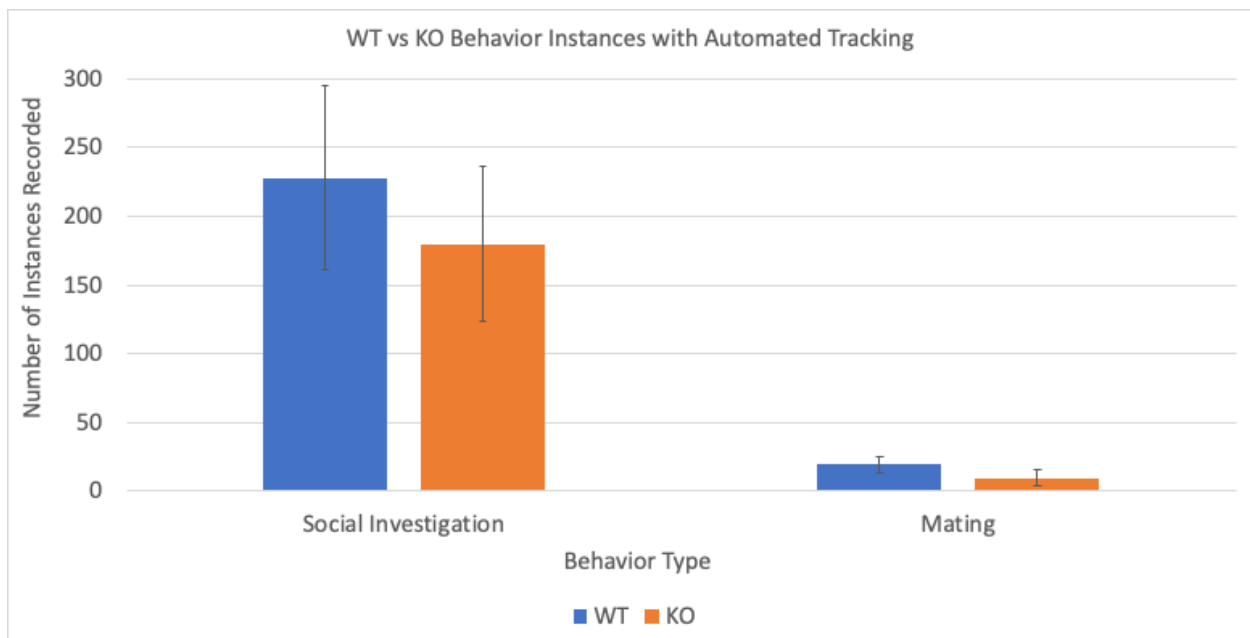
**Figure 11: Distribution of average durations of social behaviors (auto-tracking).** Each bar represents the average total time that the behavior was observed for across n=6 animals. The differences between WT and KO for both categories of behaviors are insignificant (t-test, 2-tailed, unpaired: 0.85, 0.59).



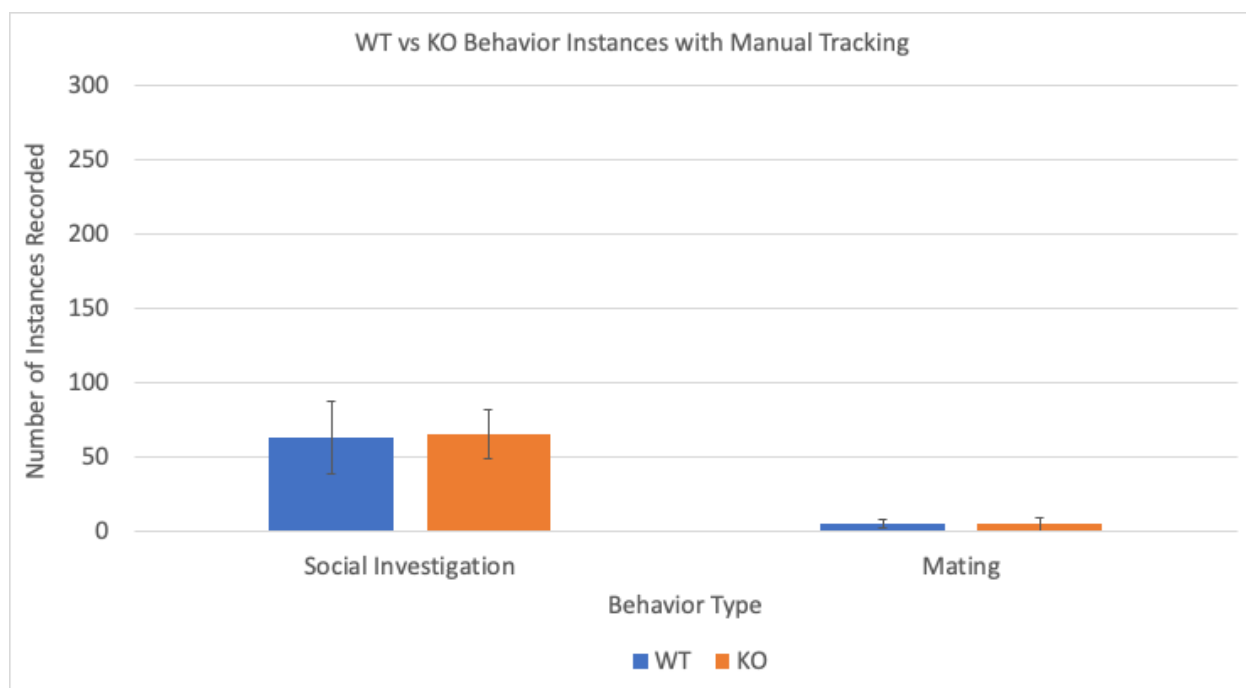
**Figure 12: Distribution of average durations of social behaviors (manual-tracking).** Each bar represents the average total time that the behavior was observed for across n=6 animals. The differences between WT and KO for both categories of behaviors are insignificant (t-test, 2-tailed, unpaired: 0.69, 0.43).



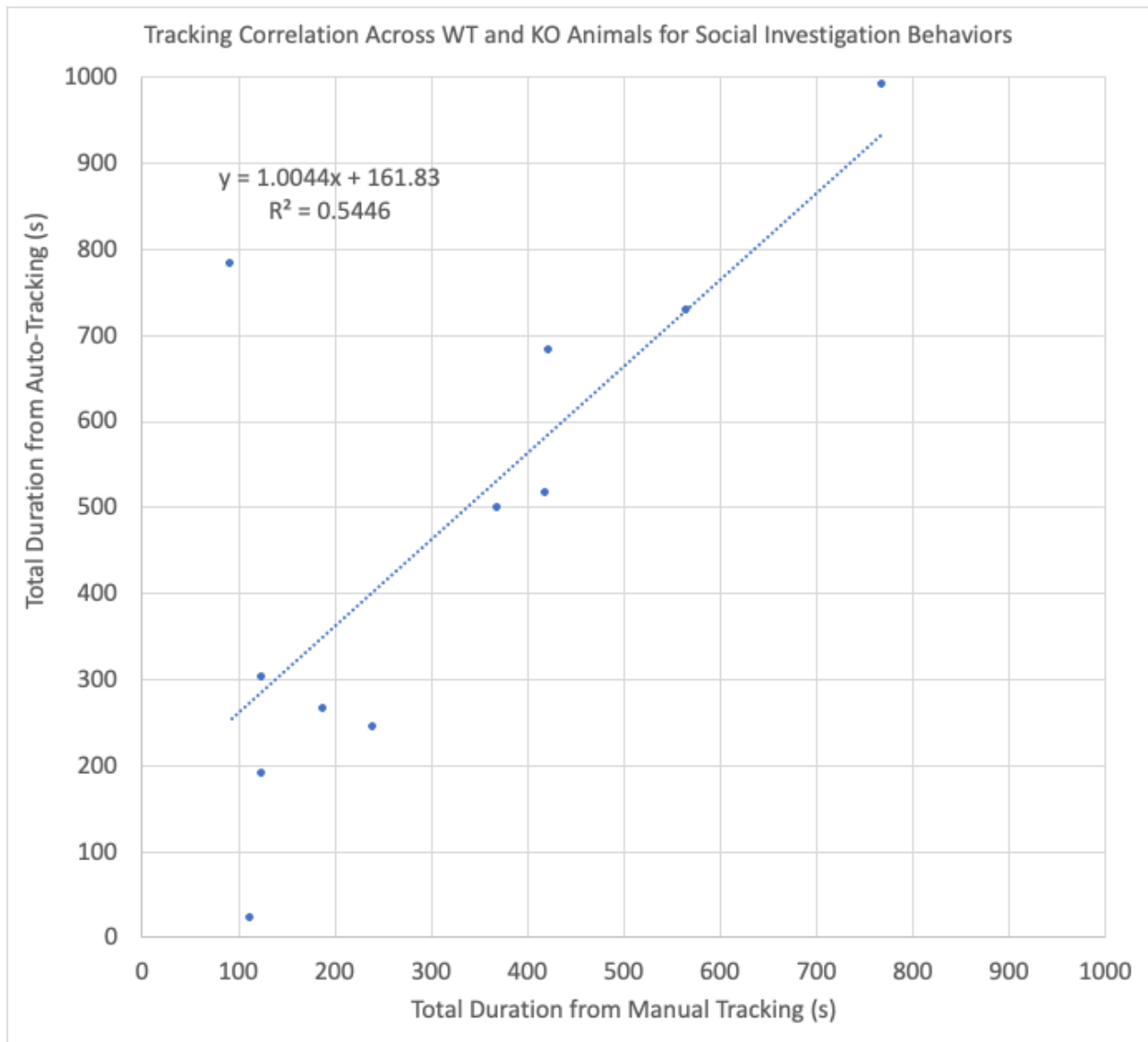
**Figure 13: Distribution of average number of instances of social behaviors (auto-tracking).** Each bar represents the average total number of instances that the behavior was observed for across  $n=6$  animals. The differences between WT and KO for all four behaviors are insignificant (t-test, 2-tailed, unpaired: 0.90, 0.39, 0.78, 0.30).



**Figure 14: Distribution of average number of instances of social behaviors (auto-tracking).** Each bar represents the average total number of instances that the behavior was observed for across  $n=6$  animals. The differences between WT and KO for both categories of behaviors are insignificant (t-test, 2-tailed, unpaired: 0.84, 0.32).

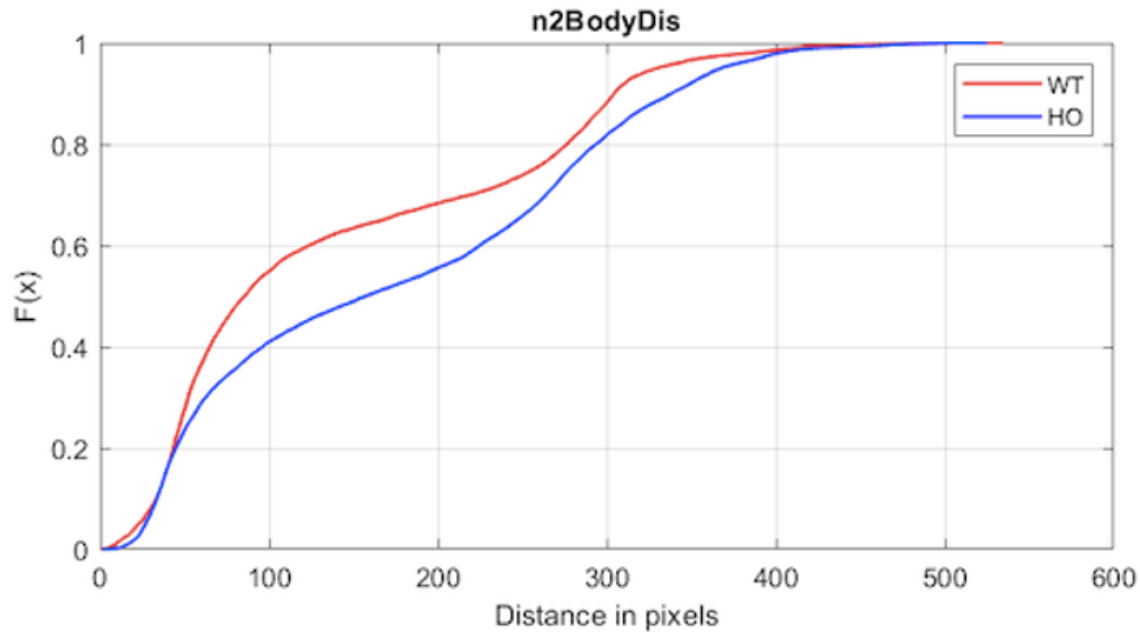


**Figure 15: Distribution of average number of instances of social behaviors (manual tracking).** Each bar represents the average total number of instances that the behavior was observed for across  $n=6$  animals. The differences between WT and KO for both categories of behaviors are insignificant (t-test, 2-tailed, unpaired: 0.88, 0.77).

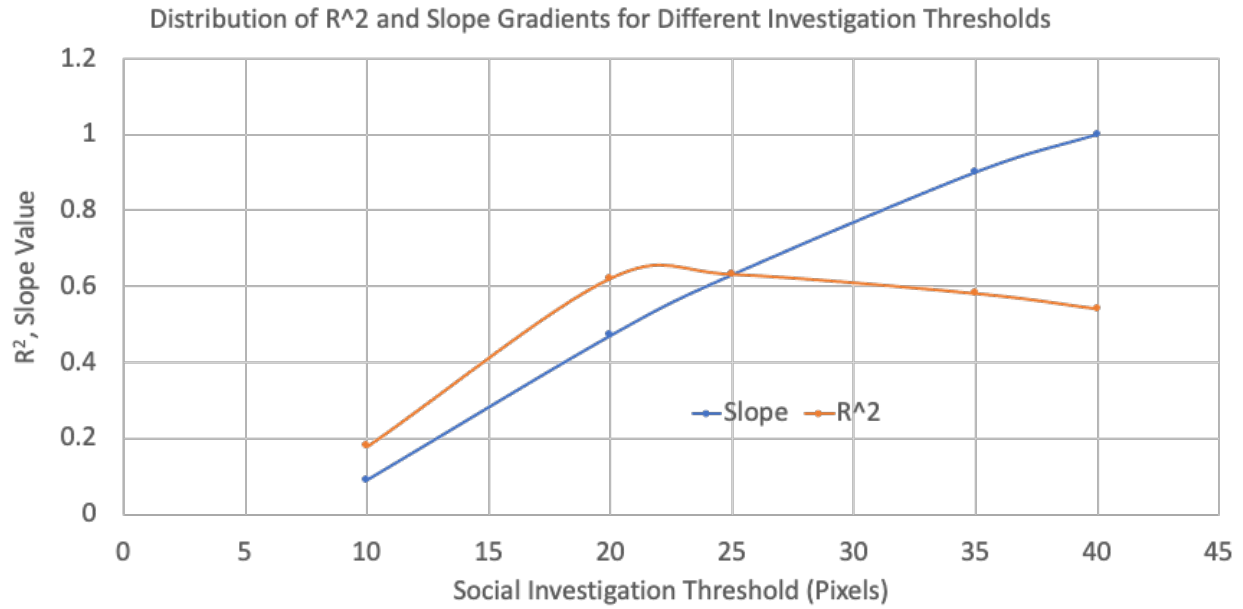


**Figure 16: Correlation between auto-tracked and manually generated results (social investigation durations).** The near-1 slope value indicates minimal difference in the total duration of social investigation captured by manual and automated tracking approaches. Each point represents a single animal (n=11).

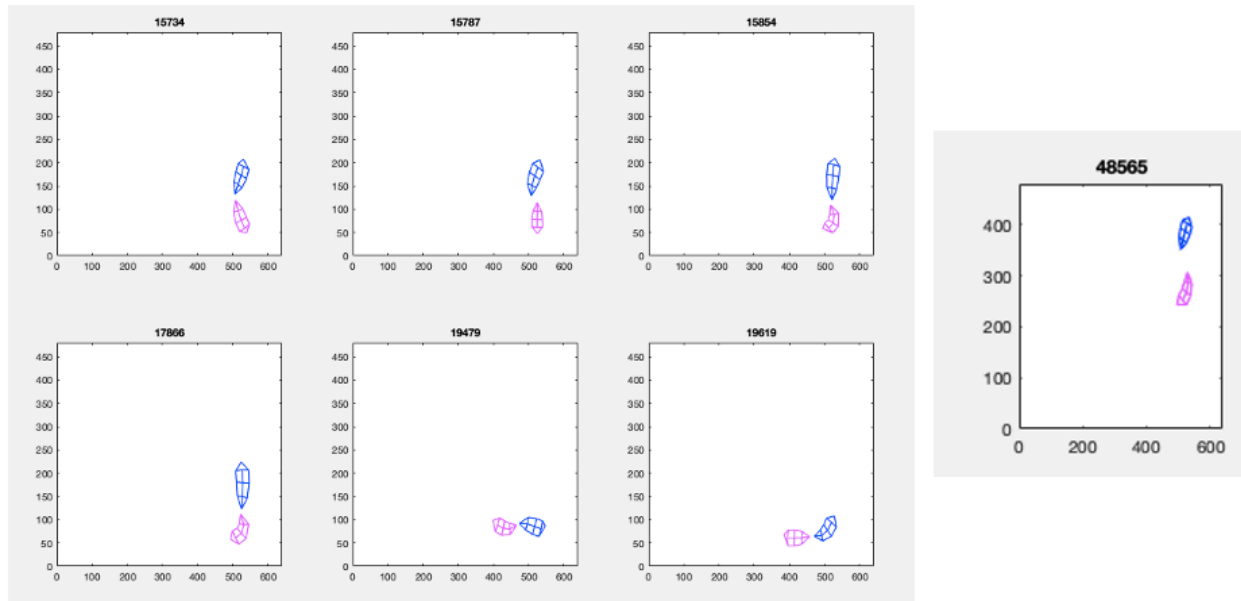




**Figure 17: Cumulative probability distribution of frames identified by auto-tracking network for various male-female distance thresholds:** The x axis illustrates a range of possible threshold values for the distance between the male nose coordinate and the female midbody coordinate. (30 Pixels = distance from vole nose to between-ears = 12 mm approx.). The y axis represents a probability density function for the percentage of frames (for either HO or WT animals) that the auto-tracking network identified to be within a given value of distance between the two animals. The results have been averaged across all animals in each group (n=5 for WT and n=6 for HO). Upon crossing a threshold of 40 pixels, until 400 pixels, WT animals on average appear to be in closer proximity to their female conspecifics than do HO animals to *their* female conspecifics.



**Figure 18: Distribution of  $R^2$  and slope values obtained for correlations between manual and auto-tracking results across all animals ( $n=11$ ):** For social investigative behavioral durations (see Methods (5) and Figures 7, 9, and 16),  $R^2$  peaks between a threshold of 20 and 25 pixels and the slope value reaches 1 at a threshold of 40 pixels.



**Figure 19: Frames identified by auto-tracking network for instances of nose-nose sniffing at a specific threshold value:** The left panel displays results that would always be considered as “true” for the behavior during manual annotation processes, given the proximity and alignment of the animals. The figure on the right represents a “false positive” result at the same threshold value. While the animals have aligned appropriately to carry out *nose-nose sniffing*, they appear to be too “distant” from each other to be considered “true” for the behavior during manual annotation processes. The results from Figure 17, however, might suggest that the frame on the right is in fact “true” for the given behavior.