
Combination of deep behavioral phenotyping with brain region and cell type specific manipulations of FKBP51

Joeri Bordes



München 2023

Combination of deep behavioral phenotyping with brain region and cell type specific manipulations of FKBP51

Joeri Bordes

A dissertation

submitted to the Faculty of Biology

at the Ludwig–Maximilians–Universität München

for the degree of

DOCTOR RERUM NATURALIUM



München 2023

First referee: PD Dr. Mathias V. Schmidt

Second referee: Prof. Dr. Laura Busse

Day of submission: 05.07.2023

Day of the oral examination: 10.11.2023

Table of contents

Table of contents	vi
Summary	ix
List of abbreviations	xi
Declaration of Contribution	xiv
1 General introduction	1
1.1 The impact of stress-related disorders	1
1.2 The terminology of stress	2
1.3 The stress response system	3
1.3.1 The SAM axis	4
1.3.2 The HPA axis	7
1.4 The psychiatric risk gene FKBP5	12
1.4.1 Clinical evidence for FKBP5 as a psychiatric risk factor	12
1.4.2 Animal studies on the effects of Fkbp5 manipulation after stress exposure	13
1.5 The Locus Coeruleus	15
1.5.1 Noradrenergic signaling in the brain: a broad pathway through the neuraxis	15
1.5.2 The role of the Locus Coeruleus in stress-related disorders	18
1.6 Rationale and thesis objectives	19
2 Advancing social behavioral neuroscience by integrating ethology and com- parative psychology methods through machine learning	21
3 DeepOF: a python package for supervised and unsupervised pattern recog- nition in mice motion tracking data	33
4 Automatically annotated motion tracking identifies a distinct social behav- ioral profile following chronic social defeat stress	39
5 Early life stress affects the acquisition of fear differently between sexes	83
6 The Fkbp5 knock-out in the noradrenergic locus coeruleus system alters social behavior in male mice	115
7 Increasing resolution in stress neurobiology: from single cells to complex group behaviors	145

8 General discussion	159
8.1 Summary	159
8.2 Next generation deep phenotyping requires more than motion tracking data .	161
8.3 Decoding the ambiguous nature of FKBP51 on psychiatric risk through brain region and cell type specific effects	164
8.4 The influence of sex on stress vulnerability: a plea for more ethologically relevant stressors	166
8.5 Closing remarks	168
Bibliography	186
Acknowledgements	190
Publications	193
Assertion/Eidesstattliche Erklärung	194

Summary

The increasing prevalence of stress-related disorders, such as major depressive disorder (MDD) has become a significant global concern, with devastating effects on individuals' personal lives and societal well-being. The exposure to severe and chronic stressors is a major risk factor for the development of such disorders, and recent traumatic events have further exacerbated this mental health crisis. The susceptibility to MDD is determined by a complex interplay of genetic, epigenetic, and environmental factors. One specific gene of significance in this context is FKBP5 (Fkbp5 in rodents), encoding the co-chaperone FK506 binding protein 51 (FKBP51). The interplay between severe stress exposure and genetic risk variants of FKBP5 has been associated with an increased vulnerability to psychopathology.

A significant symptom observed in individuals with MDD is social dysfunction, characterized by the avoidance of social interactions and the display of maladaptive behaviors, such as aggression or irritability. However, traditional preclinical assessment methods for stress-induced behavioral symptoms, such as social aversion, have faced criticism due to their reductionistic nature, often failing to capture ethologically relevant behavioral constructs. Advancements in high-throughput pose estimation tools have provided opportunities for comprehensive behavioral analysis through automatically annotated behavioral assessments. This thesis explores various tools for automatically annotated behavioral assessment in preclinical psychiatry research, employing both supervised classification and unsupervised clustering strategies.

Applying the newly established and validated deep phenotyping methods, the thesis further investigates the brain region and cell type specific role of FKBP51 across different stress models and uncovers the underlying neurobiological mechanisms and behavioral profiles using automatically annotated behavioral assessment. The effectiveness of both supervised classification and unsupervised clustering strategies is demonstrated in characterizing individual and social behavioral profiles in mice subjected to various stress conditions. Moreover, the thesis highlights the distinct sex-specific effects of different stress paradigms on the regulation of the hypothalamic-pituitary-adrenal (HPA) axis, including the expression of Fkbp5 in several

stress-related brain regions, in particular the Locus Coeruleus (LC).

Taken together, the current thesis emphasizes the importance of brain region and cell type specific regulation of Fkbp5 and underscores the benefits of automatically annotated behavioral assessment tools. This is put into perspective with future research prospects, advocating for the integration of diverse data modalities, such as in vivo measurements of stress mediators and neuronal activity recordings. This integrated approach aims to enhance our understanding of complex behaviors and the underlying molecular mechanisms. Ultimately, this can contribute to a better comprehension of the behavioral phenotypes and associated neurobiological alterations in stress-related disorders. These insights hold potential to facilitate the development of novel treatments for psychiatric disorders.

List of abbreviations

11 β -HSD	11 β -hydroxysteroid dehydrogenase
ANS	Autonomic nervous system
ASDS	Acute social defeat stress
AVP	Arginine vasopressin
BNST	Bed nucleus of stria terminalis
CeA	Central nucleus of the amygdala
Cort	Corticosterone
CRF	Corticotropin releasing factor
CSDS	Chronic social defeat stress
DDC	DOPA decarboxylase
DSM-5	Diagnostic and statistical manual of mental disorders 5th edition
DMDD	Disruptive mood dysregulation disorder
E	Epinephrine
FKBP51	FK506 binding protein 51
FKBP52	FK506 binding protein 52
Fkbp5 ^{Nat}	Fkbp5 mouse knockout line within the noradrenergic system
GCs	Glucocorticoids
GR	Glucocorticoid receptor
GRAB	GPCR activation-based
GRE	Glucocorticoid response element
GWAS	Genome-wide association studies
HPA axis	Hypothalamic-pituitary-adrenal axis
HSP	Heat shock protein
HOP	HSP70/HSP90 organizing protein
IHME	Institute for Health Metrics and Evaluation

IML	Intermediolateral cell column
KO	Knockout
LC	Locus Coeruleus
L-DOPA	L-3,4-dihydroxyphenylalanine
MDD	Major depressive disorder
MHPG	3-methoxy-4-hydroxyphenylglycol
mRNA	Messenger RNA
MR	Mineralocorticoid receptor
NE	Norepinephrine
PPIase	Peptidyl-prolyl <i>cis-trans</i> isomerase
PNS	Parasympathetic nervous system
PVN	Paraventricular nucleus of the hypothalamus
PMDD	Premenstrual dysphoric disorder
POMC	proopiomelanocortin
PTSD	Post-traumatic stress disorder
SNRIs	Serotonin and norepinephrine reuptake inhibitors
SNS	Sympathetic nervous system
SSRIs	Selective-serotonin-reuptake-inhibitors
SAM axis	Sympathetic-adreno-medullar axis
TPR	Tetratricopeptide repeat
TH	Tyrosine hydroxylase
VLM	Ventrolateral medulla
WT	Wildtype

Declaration of Contribution

The contributions to the different chapters in the current thesis, entitled ‘Investigating Brain Region-Specific Effects of FKBP51 on Stress Response - A Deep Phenotyping Approach Using Automated Behavioral Assessment Tools‘ have been done in the following way:

chapter 2

Advancing social behavioral neuroscience by integrating ethology and comparative psychology methods through machine learning. Bordes, J., Miranda, L., Müller-Myhsok, B., Schmidt, M.V.

Outline of the review: JB and MVS.

Writing of the review: JB.

Feedback on the review: JB, LM, BMM, MVS.

chapter 3

DeepOF: a python package for supervised and unsupervised pattern recognition in mice motion tracking data. Miranda, L., Bordes, J., Pütz, B., Schmidt, M.V., Müller-Myhsok, B.

LM wrote the code, with assistance of BP and BMM.

JB and MVS designed the experimental set-up.

JB and LM analyzed the data.

LM wrote the manuscript with assistance of JB.

chapter 5

Early life stress affects the acquisition of fear differently between sexes Bordes, J., Miranda, L., van Doeselaar, L., Brix, L.M., Narayan, S., Yang, H., Mitra, S., Kovarova, V., Springer, M., Müller-Myhsok, B., Schmidt, M.V.

Designing and planning the study: JB and MVS.

Conducting the experiments: JB with assistance from LvD, LMB, SN, HY, SM, VK and MS.

Analyzing the data: JB and LM. Preparing the manuscript: JB, with assistance of MVS.

chapter 6

The Fkbp5 knock-out in the noradrenergic locus coeruleus system alters social behavior in male mice Bordes, J., Bajaj, T., Chang, S., Miranda, L., Pöhlmann, M., Meccad, Y., Anderzhanova, E., Brix, L.M., van Doeselaar, L., Narayan, S., Yang, H., Mitra, S., Kovarova, V., Springer, M., de Angelis, M., Müller-Myhsok, B., Deussing, J., Gassen, N., Schmidt, M.V.

Designing and planning the study: JB and MVS.

Conducting the experiments: JB with assistance from TB, SC, MP, YM, EA, LMB, LvD, SN, HY, SM, VK, MS and Mda.

Analyzing the data: JB and LM. Preparing the manuscript: JB, with assistance of MVS.

We hereby confirm the above statements.

Joeri Bordes

PD. Dr. Mathias V. Schmidt

chapter 4

Automatically annotated motion tracking identifies a distinct social behavioral profile following chronic social defeat stress

Bordes, J.*, Miranda, L.*, Reinhardt, M., Narayan, S., Hartmann, J., Newman, E.L., Brix, L.M., van Doeselaar, L., Engelhardt, C., Dillmann, L., Mitra, S., Ressler, K.J., Pütz, B., Agakov, F., Müller-Myhsok, B., Schmidt, M.V.

*Shared authorship

Originally published in: *Nature Communications*

Contributions:

JB and MVS conceived the study.

LM wrote the DeepOF module, with primary technical assessment from FA, BP, and BMM.

JB and MR performed the experiments.

JB and LM analyzed the data and wrote the manuscript.

We hereby confirm the above statements.

Joeri Bordes

Lucas Miranda

PD. Dr. Mathias V. Schmidt

Chapter 1

General introduction

1.1 The impact of stress-related disorders

The number of people suffering from stress-related disorders, such as major depressive disorder (MDD), and anxiety-related disorders, including post-traumatic stress disorder (PTSD), has rapidly increased in the last decades [1]. In 2019, the Institute for Health Metrics and Evaluation (IHME) published the Global Burden of Disease study, which indicated that depression and anxiety-related disorders have shown an enormous increase in prevalence in the last decades and have among the highest prevalence of all mental disorders [2]. This is why disorders, such as MDD, have been described by the World Health Organization as the leading cause of worldwide disability with an estimated 280 million people suffering from MDD in 2019 [1, 2]. These disorders have a devastating effect on the patients' personal life impacting their social functioning, which can lead to social withdrawal from society and in severe cases to suicide [3, 4]. Furthermore, there are also severe consequences for society related to a high economic burden due to both direct costs of healthcare and indirect costs via the loss of productivity and income [5]. A major risk factor for the development of such disorders is the exposure to stressful events that are perceived as particularly severe and/or chronic [6–9]. Recent traumatic events around the world, for instance, pandemics, wars, and natural disasters have had a detrimental impact on the mental health of the population and predictions show an even further increase in the number of people suffering from these stress-related disorders [10–12]. The treatment of stress-related disorders remains difficult with regard to efficacy and specificity. In MDD, the most common line of pharmacological treatment relies on drugs influencing the monoaminergic system in the brain via selective serotonin reuptake inhibitors (SSRIs), and serotonin and norepinephrine reuptake inhibitors (SNRIs). However, the efficacy of these drugs has remained low, as the response rates are at best around 53.8% [13]. In addition, the severe and high number of side effects can even further hamper daily life activities

and often result in patients quitting medications without improvements in their symptoms [14–16]. One of the explanations for the ineffectiveness of treatment is the high comorbidity of MDD with other psychiatric disorders, which makes the classification extremely difficult [17, 18]. Unsurprisingly, MDD has a high heterogeneity between the experienced symptoms of patients [19]. In 2013, the American Psychiatric Association published the fifth and latest version of the Diagnostic and statistical manual of mental disorders (DSM-5) including further specifications between different types of depression, in which a separate classification was made for disruptive mood dysregulation disorder (DMDD) and premenstrual dysphoric disorder (PMDD) [20, 21]. However, heterogeneity of MDD symptoms remains a problem, and important novel distinctions between MDD patients are continuously discovered [22]. The lack of understanding of the neurobiological underpinnings of MDD is problematic and needs to be addressed. An increased understanding of the neurobiological mechanism allows for a more precise screening of potential novel anti-depressant drug treatments that could target different neurobiological mechanisms, which ultimately could lead to more personalized drug treatments that show a higher efficacy and have reduced side effects.

1.2 The terminology of stress

The terminology "stress" is commonly used in our daily life to describe our experiences or emotional state. However, there is a strong negative association with stress experience, which in certain cases can be appropriate since stress can lead to severe negative consequences, including increased vulnerability towards stress-related disorders. The exposure to stress does not always lead to a negative outcome and via adaptive processes can have a positive physiological and behavioral outcome, for instance, on performance, motivation, and mood [23–26]. The history of stress research dates back to 1872 when Claude Bernard described the first concept of the stability of an internal environment within the body (*milieu intérieur*) [27]. This concept was then further conceptualized by Walter Cannon, who introduced the term "homeostasis", which is the reaction of the body to constantly adapt in order to maintain a physiological equilibrium [28]. The fundamental basis for the definition of our current understanding of stress was solidified by Hans Selye in 1936 [29], who defined stress as the nonspecific response of the body to both positive, as well as negative demands. The body is constantly adapting to maintain a physiological equilibrium (homeostasis) and the exposure to specific internal or

external stimuli (stressors) threatens the homeostasis of the body. However, this definition of stress was seen as too broad, as it included any stimulus that evokes a physiological reaction. Therefore, McEwen and Stellar in 1993 [30] introduced the concept of allostasis, which refers to the process by which stressors (whether physiological or psychological) are mediated in the body via hormonal mediators (such as glucocorticoids and catecholamines) to create stability through change, regulating adaptation, homeostasis maintenance, and survival to stress [23, 31, 32]. Importantly, prolonged and severe exposure to stressors can lead to a build-up of allostatic load, which changes the properties of the stress response system and increases the vulnerability towards stress-related disorders [23, 32]. Even though the initial definition of stress has been around for over 80 years, there is still an ongoing discussion about the exact description of stress. The understanding of stress and its definition is changing over time, which is influenced by the increased attention to mental disorders within society. A recent review from Richter-Levin & Sandi [33] highlights the increased association of stress exposure with the development of psychopathology, even though the more common reaction towards stress exposure is adaptation and building resiliency towards stress. Therefore, a distinction is proposed between the terminology of stress and stressful experiences, which refers to adaptation and resiliency towards stress, whereas the term "trauma", refers to exposures that lead to negative consequences, such as psychopathology. It is important to recognize that the definition of stress has a significant impact on society and ongoing discussion about its definition is crucial to reflect the current societal understanding.

1.3 The stress response system

The presence of a potential or actual threat, as represented by stress, requires immediate and future adaptation in the physiology and behavior of the individual. The stress response system is an evolutionarily conserved model that activates different neuroendocrine pathways in parallel, consisting of a complex symphony of many different mediators, including neurotransmitters, peptides, and steroid hormones. This complexity of the stress response system is necessary to appropriately respond to the many different factors influencing the stress response including, the severity, predictability, type (psychological or physical), and duration (acute or chronic) of the stress exposure, but also the age (early, adult, or old age), and the sex of the animal (female or male) [34, 35]. The major stress response axes consist of the fast-acting

(within seconds) sympathetic-adreno-medullar (SAM) axis via catecholaminergic signaling, and the slower-acting (minutes to hours, up to days) hypothalamic-pituitary-adrenal (HPA) axis via steroid hormone signaling. Both systems have direct and indirect influences on the periphery of the body by restoring homeostasis through different processes, such as energy mobilization, but also influence the central nervous system via processes such as learning, memory, emotional responses, and decision-making [24, 34, 36]. Even though these systems are activated rapidly, they can have long-lasting effects in a timescale from seconds to many years.

1.3.1 The SAM axis

In response to a stressful event or threat, the body needs to adapt rapidly in order to deal with the issue at hand. This rapid stress response is initiated via the autonomic nervous system (ANS), which regulates many different bodily processes, such as heart rate and pupil dilation, but also the activity of the organs, such as the inhibition or stimulation of the digestive tract [37, 38]. The ANS is subdivided into the sympathetic nervous system (SNS), which is linked to the SAM axis, which initiates the "fight or flight" system, and the parasympathetic nervous system (PNS), which is linked to the "rest and digest" system. These systems have opposite actions in the majority of their functions, where one system stimulates and the other inhibits a certain physiological response [39]. The ANS activates the SNS and inhibits the PNS in response to stressful stimuli, which, among many other physiological processes, increases energy expenditure and inhibits the digestive system.

The fast processing of a stressor via the SAM axis is orchestrated by multiple brain regions in the brainstem and the hypothalamus, which include the Locus Coeruleus (LC), the paraventricular nucleus of the hypothalamus (PVN), and the ventrolateral medulla (VLM). These brain regions can directly innervate the intermediolateral cell column (IML), a group of preganglionic neurons of the SNS [36, 39, 40], see Figure 1.1. These preganglionic neurons of the IML lie within the lateral gray column of the spinal cord and are relatively short. They use acetylcholine to activate nicotine receptors to communicate to the post-ganglionic neurons, which in turn are much longer and travel to the peripheral effector sites, including the different organs and tissues sites of the body [37, 41].

The majority of the post-ganglionic neurons release NE and act on various adrenergic

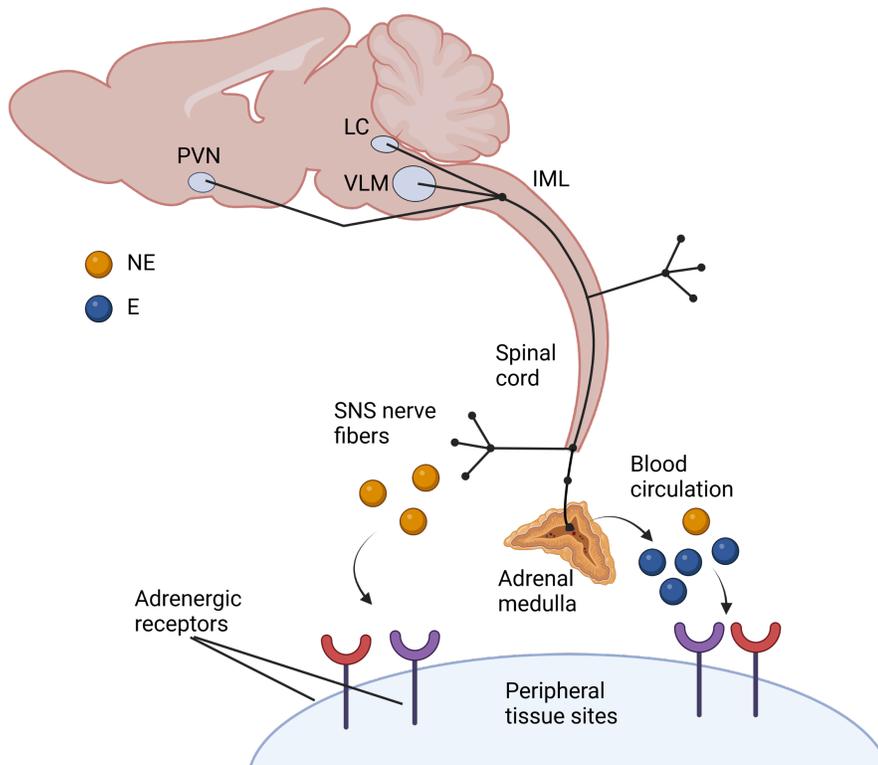


Figure 1.1: **The SAM axis stress response system.** In response to different types of stressors, the SAM axis is rapidly activated and shows a fast effect on the brain and body. This system is also called the "fight or flight" response mechanism due to its major role in preparing the body for immediate action against threats, by increasing the heart rate and blood pressure, pupil dilation, and blood flow to skeletal muscles, while decreasing blood flow to the digestive tract. The SAM axis is initiated via multiple brain regions in the brainstem and the hypothalamus, including, the locus coeruleus (LC), paraventricular nucleus of the hypothalamus (PVN), and the ventrolateral medulla (VLM), which directly innervate the intermediolateral cell column (IML). The IML communicates to the post-ganglionic neurons, which travel to the peripheral effector sites, including the different organs and tissue sites of the body. The adrenal medulla is one of the nodes that is innervated by the SNS and in turn releases epinephrine (E) and norepinephrine (NE) from the adrenal medulla into the bloodstream. Via the bloodstream E and NE can reach peripheral organs and tissues in seconds to activate the fight or flight system via further stimulation or inhibition of the specific bodily process. Adapted from [42]

receptors (α 1-2 and β 1-2 receptors) that depending on the effector site can have specific inhibitory or excitatory effects [41]. The adrenal medulla is one of the nodes that is innervated by the SNS, and in turn releases E (around 80% of the output), but also NE from chromaffin cells in the adrenal medulla into the bloodstream. Via the bloodstream E and NE can reach peripheral organs and tissues in seconds to activate the fight or flight system via further stimulation or inhibition of specific bodily processes, such as the increase of heart rate and blood pressure, and blood flow to skeletal muscles, while decreased blood flow to the digestive tract [36, 39, 43, 44]. In addition, some bodily processes are directly mediated via noradrenergic projections from the LC, for instance for pupil dilation [45, 46].

Both E and NE, are considered catecholamines, as well as monoamines, as they contain a catechol group (a benzene ring with two hydroxyl side groups at 3- and 4-positions) and a single side-chain amine[47]. The synthesis pathway of catecholamines (see Figure 1.2) starts with the conversion of the amino acid L-tyrosine to L-3,4-dihydroxyphenylalanine (L-DOPA) by the rate-limiting enzyme tyrosine hydroxylase (TH). Then, L-DOPA is decarboxylated into dopamine by DOPA decarboxylase (DDC), which in turn is converted into norepinephrine by dopamine β -hydroxylase. Norepinephrine is converted by phenylethanolamine-N-methyltransferase into epinephrine [48].

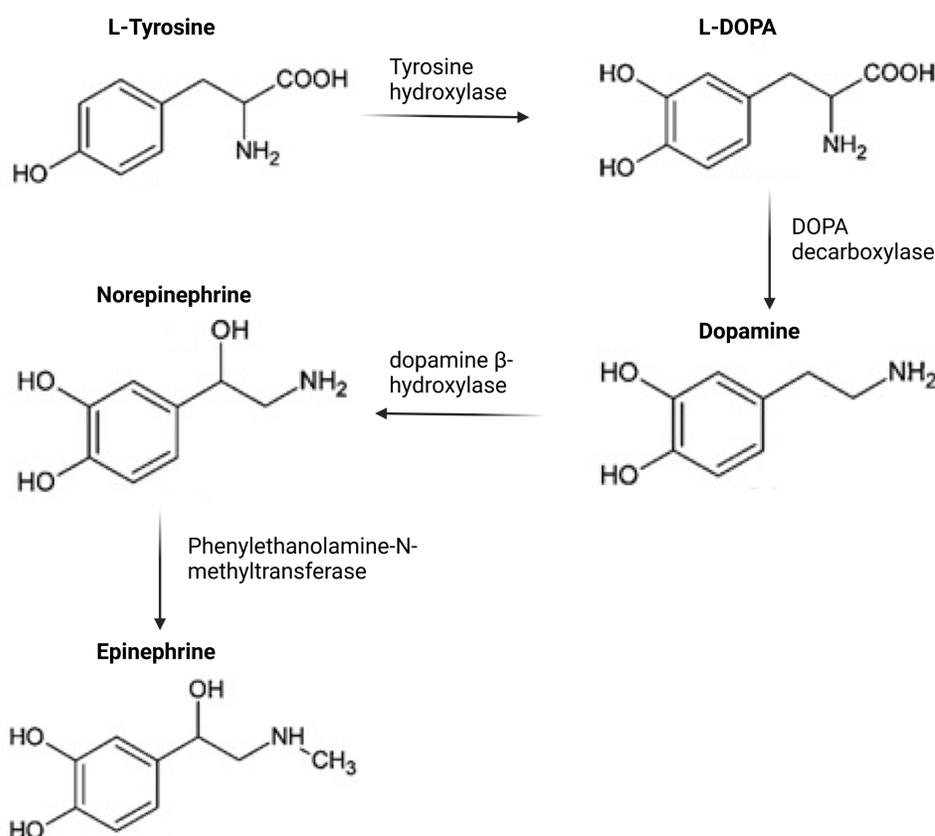


Figure 1.2: **The synthesis and metabolism of catecholamines.** Adapted from [48]

The effects of the hormonal release of NE and E from the adrenal medulla are peripheral since these catecholamines cannot pass the blood-brain barrier and have therefore no effect in the brain [49]. However, a separate neurocircuitry fight or flight response mechanism is in place via the broad signaling pathways of the monoaminergic neurocircuitry, including NE, dopamine, and serotonin [34, 50–52]. There are many factors that determine the degree to

which and to what extent the monoaminergic systems are activated, such as the type (e.g. physical or psychological), the severity (acute or chronic), and the controllability (unpredictable or predictable) of the stressor [34, 36, 50]. The different monoaminergic neuronal circuits regulate specific immediate behavioral adaptations for instance, NE signaling has been linked to increasing vigilance and attention bias, while dopamine signaling can increase risk assessment processes, and serotonin signaling is involved in post-stress anxiety regulation [52, 53].

1.3.2 The HPA axis

In response to a stressful event, the HPA axis is activated in parallel with the "fight or flight response", but has a slower and prolonged stress response mechanism. The HPA axis is primarily activated during stressful events, but other activities such as exercise, intercourse, or changes in appetitive reward schedules have also been linked to an activation of the HPA axis [54, 55]. The neuroanatomical circuitry of the HPA axis consists of the activation of different neuronal populations in the PVN via direct and indirect neuronal input from a number of brain regions related to stress evaluation, including cortical and limbic regions, but interestingly is also innervated via the NE signaling pathway coming from the brain stem structure, nucleus of the solitary tract [56]. The main PVN stress-related neuronal networks are the hypophysiotropic neurons producing corticotropin-releasing factor (CRF) and the arginine vasopressin (AVP) expressing neurons, which are located in the parvocellular subdivision of the PVN and project into the circulation of the median eminence [40, 57]. The release of CRF and AVP then act via CRF-1 and VP-1b receptors, respectively, in the endocrine cells of the anterior pituitary (corticotrophs). This activates the rapid cleavage of the proopiomelanocortin (POMC) hormone into the adrenocorticotrophic hormone (ACTH), which is then released via secretory vesicles into the bloodstream [58, 59]. The exocytosis of ACTH is primarily controlled via CRF acting at the CRF-1 receptor side, but can also be stimulated via the AVP-VP-1b pathway [60]. Ultimately, the release of ACTH into the bloodstream triggers the synthesis and diffusion of glucocorticoids (GCs), cortisol in humans, and corticosterone (Cort) in rodents, by cells in the zona fasciculata layer of the adrenal cortex via the melanocortin 2 receptor, see Figure 1.3. Increased levels of Cort in the bloodstream typically can be found only 3-5 minutes after the onset of the HPA axis, whereas ACTH is

more rapidly released into the circulation. This is due to the necessary time to synthesize Cort via a series of enzyme-mediated reactions from cholesterol, and subsequently to diffuse Cort into the circulation [60, 61]. Cort levels typically peak within 30 minutes after acute stress exposure and are usually back to baseline after 60-90 minutes [62].

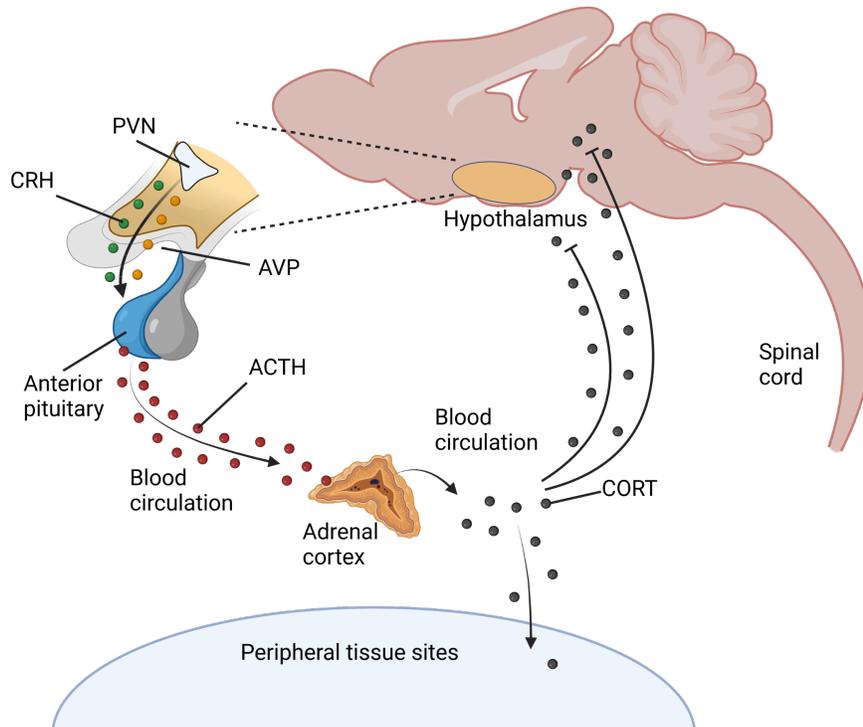


Figure 1.3: **The HPA axis stress response system.** In response to different types of stressors, the HPA axis is activated and shows a slower but more sustained effect on the brain and body compared to the rapid actions of the SAM axis. The HPA axis is initiated via the release of corticotropin-releasing factor (CRF) and arginine vasopressin (AVP) from different neuronal populations of the paraventricular nucleus of the hypothalamus (PVN), which in turn triggers the release of adrenocorticotropic hormone (ACTH) from the anterior pituitary into the bloodstream. Ultimately, the release of ACTH into the bloodstream triggers the synthesis and diffusion of GCs from the adrenal cortex, exerting its function in the periphery as well as in the brain. After the initial activation of the HPA axis, different feedback loops are in place to appropriately regulate the cognitive response and the HPA axis activity, which is mediated via Cort-GR/MR binding in several brain regions, such as in the hypothalamus, hippocampus, and frontal regions.

The basal activity of the HPA axis is influenced by both ultradian and circadian rhythms, which lead to pulsatility of basal Cort secretion within the hourly and daily time windows [63]. In rodents, this leads to an increase of baseline Cort several hours before lights-off, with a daily peak at around the lights-off moment. [64]. The timing of stressors is crucial, as these pulses in Cort secretion influence the response of the HPA axis to acute stress. A stressor that occurs during the rising phase of an ultradian Cort pulse rather than the falling phase has been implicated with a larger stress-induced Cort response [65]. The peripheral effects of

the HPA axis are partially complementary to the fight or flight response system, but also have specific additional effects. The peripheral effects of the HPA axis include the regulation of metabolism, immune function, inflammation, cardiovascular function, and electrolyte balance. The lipid-soluble nature of the Cort molecule enables it to easily pass through the blood-brain barrier, regulating a variety of biological processes, not just in the body, but also in the brain, which is not possible for the fight or flight hormones E and NE [49].

Cort exerts its effects in various cells and tissues throughout the body and brain by binding to two different receptor types, the glucocorticoid receptor (GR) encoded by the NR3C1 gene, and the mineralocorticoid receptor (MR) encoded by the NR3C2 gene. The binding affinity of these receptors varies; the MR has a 10-fold higher affinity to Cort compared to the GR. Therefore, at baseline conditions with low levels of circulating Cort, there is a high occupation of the MR, while upon activation of the HPA axis and the increase of circulating Cort, not only the MR but also the GR will be occupied by Cort [43]. The GR is present in almost all tissues in the body, including the nervous-, immune-, cardiovascular-, respiratory-, reproductive-, musculoskeletal, and integumentary systems, and is also widely expressed throughout the brain, with the highest density in the anterior pituitary, PVN, and limbic-prefrontocortical regions [66]. Meanwhile, peripheral MR is specifically found in epithelial cells of the kidney, bladder, and intestines but also in other cell types and tissues of among others, the heart. MR expression in the brain follows a specific expression pattern, with the highest density in the limbic regions, such as the hippocampus and amygdala [66, 67]. These steroid receptors are named after their functions in the periphery, where the MR is important for the mineral balance and the GR for glucose metabolism [68]. More specifically, one of the functional roles of the MR in the periphery is to regulate ion and fluid balance via sodium homeostasis in epithelial cells. In order to distinguish this function from its Cort binding functions, there is the local inactivation of Cort by 11 β -hydroxysteroid dehydrogenase (11 β -HSD) type 2 to 11-deoxycorticosterone (cortisone in humans), which cannot bind steroid receptors. Therefore, in those peripheral areas, there is MR binding to the lower circulating concentrations of the other endogenous ligand of the MR, aldosterone, that is involved in sodium homeostasis[69]. Importantly, in the brain Cort's binding properties are reactivated via the 11 β -HSD type 1 enzyme, which means that the neuronal effects of the MR are driven by Cort-MR binding [70].

Cort has an important role in the brain to appropriately respond to stressful events but is restricted to the neuronal circuits that express MR and/or GR. An important function of Cort in the brain is to regulate the activity of the HPA axis via a negative feedback loop by binding to the GR at the PVN and the anterior pituitary sides, which shuts off the release of CRF and ACTH, respectively [71, 72]. In addition, Cort has specific stress-related brain region effects, which influence the cognitive, emotional, and neuroendocrine processing via both GR and MR binding, such as in the hippocampus, LC, nucleus of the solitary tract, and the central amygdala, but also via GR binding alone, such as in the dorsal raphe nucleus, medial amygdala, basolateral amygdala, and the medial prefrontal cortex [34], see (Figure 1.4).

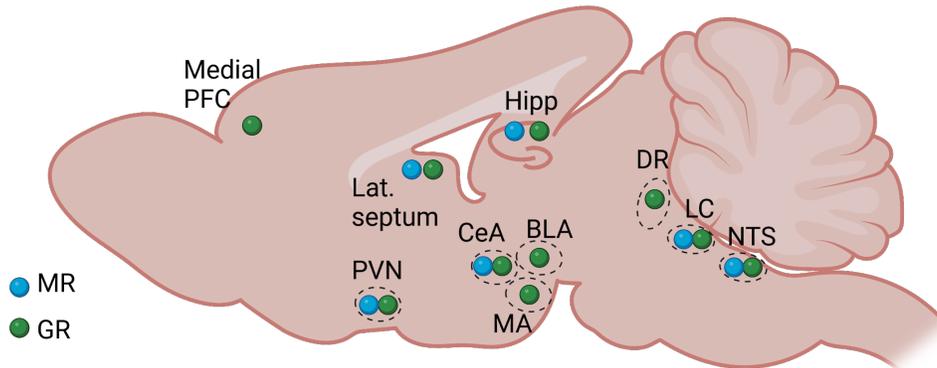


Figure 1.4: **The expression of Glucocorticoid receptor (GR) and Mineralocorticoid receptor (MR) throughout the brain.** The role of GCs in the brain is restricted to the neuronal circuits that express MR and/or GR. These brain regions are strategic hubs that connect networks involved in diverse aspects of the brain's stress response. An important function of Cort in the brain is to regulate the activity of the HPA axis via a negative feedback loop by binding to the GR at the paraventricular nucleus of the hypothalamus (PVN) and the anterior pituitary sides, which shuts off the release of corticotropin-releasing factor (CRF) and adrenocorticotropic hormone (ACTH), respectively. In addition, Cort has specific stress-related brain region effects which influence the cognitive, emotional, and neuroendocrine processing via both GR and MR binding, such as in the hippocampus (Hipp), locus coeruleus (LC), nucleus of the solitary tract (NTS), and the central amygdala (CeA), but also via GR binding alone, such as in the dorsal raphe nucleus (DR), medial amygdala (MA), basolateral amygdala (BLA), and the medial prefrontal cortex (medial PFC). Adapted from [34]

The MR- and GR-mediated processes can be complementary, but also opposing however, their molecular signaling pathways have a similar structure, as they are nuclear receptors that are able to affect gene transcription via the Cort-MR/GR complex [73]. These receptors for example, the GR can then transfer to the nucleus and bind to the glucocorticoid response element (GRE) in the regulatory regions of target genes (genomic actions), but also can have their effects at the cell membrane via rapid GRE-independent interactions (non-genomic actions) [74, 75]. In the absence of Cort or other ligands, the GR receptor remains in the

cytoplasm and is complexed to chaperone proteins via several heat shock proteins (HSP) and their additional machinery. The conformation of the GR-chaperone complex structures folds the GR into a higher or lower affinity for hormone binding states, which in turn tightly regulate the GR translocation to the nucleus [73, 76]. The GR receptor, in reticulocyte lysate, is bound by HSP70 and HSP40, which facilitate general protein folding processes. Then, the co-chaperoning protein, HSP70/HSP90 organizing protein (HOP) binds independently to HSP90 via the tetratricopeptide repeat (TPR) domain and functions as an adaptor of HSP90 to the HSP70-GR complex, see Figure 1.5. The next conformational change includes the competitive replacement of HOP by other co-chaperones that bind via the TPR domain on HSP90, including the FK506 binding proteins, in which either the immunophilin FK506 binding protein 51 (FKBP51) or FK506 binding protein 52 (FKBP52) is introduced into the conformation. This structure is stabilized by the ATP-dependent association of HSP90 with p23. In this conformation, HSP90 can influence the receptor ligand-binding domain, in which it can promote and stabilize a conformational change that establishes high-affinity Cort binding [76–79]. The immunophilins, FKBP51 and FKBP52 are conserved proteins that bind immuno-suppressant drugs, such as FK506 and rapamycin. The FKBP51 and FKBP52 proteins contain two FK-domains and three TPR motifs [80] and are highly homologous proteins with a 54.7% identity and 88.1% similarity in their amino acid sequences [81]. Despite the strong similarity between the two FKBP5s, they have opposing roles in the context of GR signaling. The GR:FKBP51 conformation stabilizes the GR structure, but decreases the binding to Cort and slows down the nuclear translocation of the GR complex. In contrast, the GR:FKBP52 conformation promotes a high-affinity Cort-binding state via the PPI domain binding. In addition, FKBP52 interacts with the motor protein dynein and promotes efficient nuclear targeting of GR [82, 83]. The promotion of nuclear gene transcription at the GRE sites via GR:FKBP52 initiates the transcription of several GR-responsive genes, which includes the initiation of an ultra-short negative feedback loop via the upregulation of FKBP5, the gene encoding the FKBP51 protein decreasing the Cort-GR binding sensitivity and nuclear translocation [84, 85].

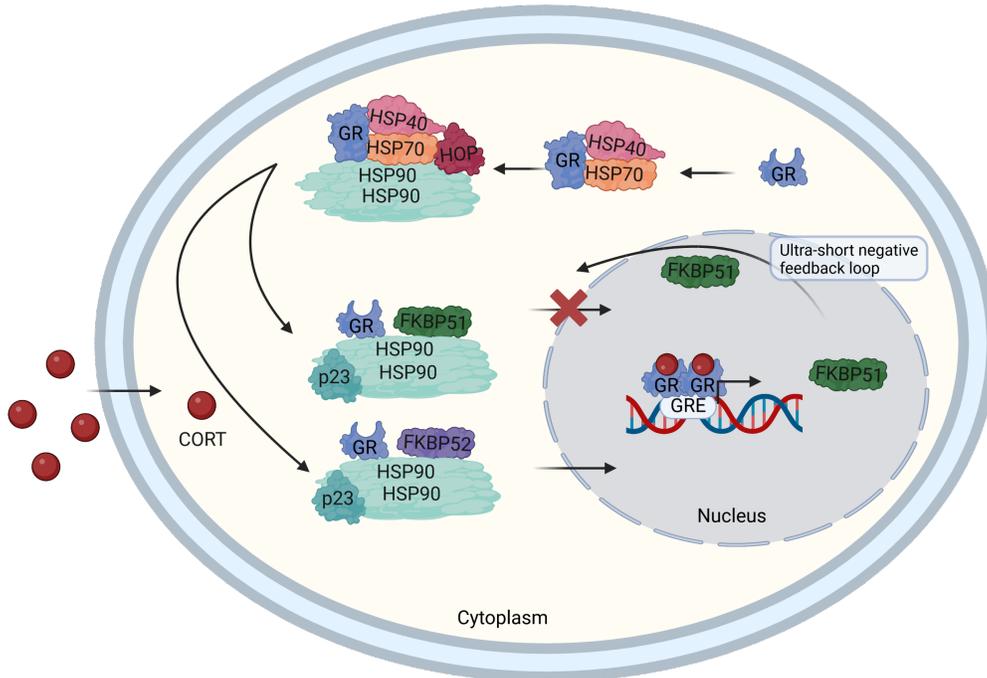


Figure 1.5: **Schematic illustration of the cellular activation of the glucocorticoid receptor (GR) nuclear translocation** After HPA axis activation, the increase in free circulating GCs can enter through the plasma membrane of the cell. In the cytoplasm of the cell, the GR complex is built, starting with the binding of HSP70:HSP40 to the GR that further unfolds the GR. This complex can then be bound via HOP which allows HSP90 to bind to the GR:HSP70:HSP40 complex. Subsequently, different heterocomplexes can be built via the binding of cochaperones, including FKBP51 and FKBP52, which bind to HSP90 and thereby release HOP. The p23 is needed to stabilize the GR-complex with HSP90 and the cochaperone. The FKBP51 conformation decreases the binding to Cort and slows down the nuclear translocation of the GR complex, whereas the FKBP52 conformation promotes a high-affinity Cort-binding state and allows the nuclear translocation of the GR complex. The translocation of the GR-complex to the nucleus allows the binding towards glucocorticoid response elements (GREs), which in turn initiates the transcription of several GR-responsive genes, which includes the initiation of an ultra-short negative feedback loop via the upregulation of FKBP5, the gene encoding the FKBP51 protein, decreasing the Cort-GR binding sensitivity and nuclear translocation. Adapted from [79]

1.4 The psychiatric risk gene FKBP5

1.4.1 Clinical evidence for FKBP5 as a psychiatric risk factor

The response to acute stress is important for the body to maintain homeostasis, but when stress becomes chronic, it can build up to an allostatic load, which is a risk factor for developing psychiatric disorders [86]. The vulnerability towards stress-related disorders, such as MDD and PTSD is determined by the combined effect of genetics, epigenetics, and the environment [43, 87]. Environmental factors in the form of severe and/or chronic stress exposure (such as, childhood maltreatment) can lead to a disturbed HPA axis system. This is heavily influenced by the dysfunction of the GR-complex, in which the negative feedback system is impaired and the HPA axis becomes hyperactive, as has been observed in the increased

baseline levels of cortisol in MDD patients [88, 89]. In particular, the heritability of MDD has been estimated at 37% using a meta-analysis of twin studies [90], and 32% using genomic similarity analysis of SNPs between unrelated patient and control individuals [91], indicating a crucial genetic role. Human genetic studies have identified different genomic variations that are involved in MDD, such as NR3C1, NR3C2, 5-HTT, SLC6A15, AVP, and FKBP5 [92–97]. In addition, the use of genome-wide association studies (GWAS) has allowed the identification of a large number of novel risk loci related to MDD [98, 99].

The FKBP5 gene has been identified with several polymorphisms that alter the gene methylation and expression and increase the vulnerability towards stress-related disorders [97, 100–106]. In particular, the risk allele of the rs1360780 leads to a higher FKBP51 protein expression, which alters the GR-signaling pathway and has been associated with an increased recurrence of depressive episodes in MDD patients [97]. Moreover, altered FKBP5 expression has been linked to abnormal HPA axis negative feedback signaling in MDD patients [107]. The combination of environmental challenges, such as childhood maltreatment, and the genetic predisposition of the FKBP5 gene has been found to lead to specific DNA demethylation in the GRE parts of the FKBP5 gene, resulting in an increased stress-driven FKBP5 expression and long-term dysregulation of the cortisol-GR signaling pathway, resulting in an increased risk for the development of stress-related disorders [100, 101]. These studies showed that the manipulation of FKBP5 can have severe consequences, which led to the question if the manipulation of the FKBP5 gene can also favor a protective effect towards the vulnerability of stress-related disorders. The genetic manipulation of FKBP5 is difficult to assess in humans but has yielded important findings using genetic mouse models.

1.4.2 Animal studies on the effects of *Fkbp5* manipulation after stress exposure

FKBP51 in mice seems to have a similar molecular role as observed in humans, in which it is involved as a co-chaperone protein of the HSP90 in binding the GR complex. Mouse *Fkbp5* mRNA is approximately 85% identical to the human FKBP5 gene, and even 87% between the FKBP51 proteins [108]. The high similarity of the GR signaling pathway between mice and humans has enabled a more in-depth mechanistic analysis of FKBP51. The expression pattern of *Fkbp5* mRNA in mice is often, but not exclusively, expressed in regions

of GR expression. A high expression of the *Fkbp5* gene was observed in, among others, the hypothalamus, amygdala, hippocampus, bed nucleus of stria terminalis, and LC [109], see Figure 1.6. In addition, the exposure to different acute stressors showed specific increases of *Fkbp5* expression, with the PVN and central amygdala responding to restraint stress, and the hippocampus to food deprivation [109].

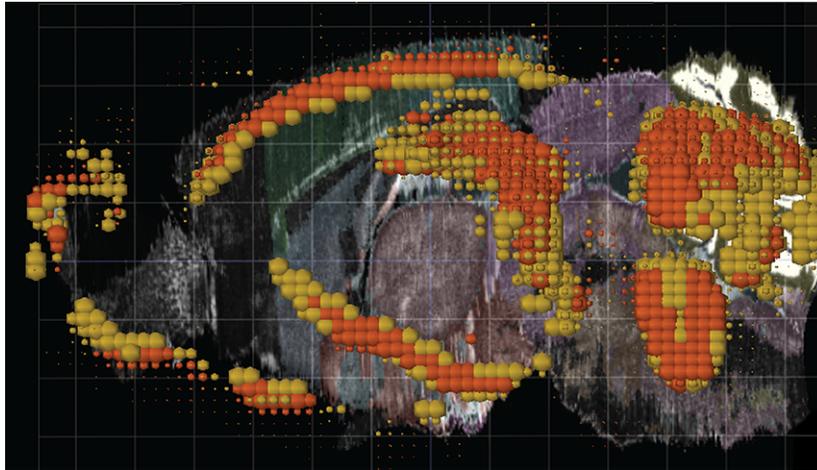


Figure 1.6: **The expression pattern of *Fkbp5* throughout the mouse brain.** *Fkbp5* expression in the mouse brain has specific expression hotspots in several stress-related brain regions. High expression of *Fkbp5* is observed in among others, the hypothalamus, amygdala, hippocampus, bed nucleus of stria terminalis, and Locus Coeruleus. Image is generated using the Allen Brain atlas 3D image viewer [110].

The exposure to chronic stress also increased the expression of *Fkbp5* throughout various stress-related brain regions, including the nucleus accumbens, hippocampus, amygdala, and prefrontal cortex [111, 112]. Mouse genetic knockout (KO) studies, in which transgenic animals are lacking FKBP51 expression have no negative consequences on their life expectancy, glucose tolerance, blood composition, and cytokine profiles [113, 114]. In addition, FKBP51-KO mice do not show behavioral alterations in exploratory drive, locomotor activity, anxiety-related behavior, or cognition under baseline conditions [115, 116]. However, exposure to different acute and chronic stress models revealed a resilient phenotype in FKBP51-KO animals. More specifically, following acute stress exposure a reduction in Cort level response was observed. In addition, more active stress-coping behavior [115, 117], but also enhanced cognitive flexibility [113], and a healthier sleeping pattern [118] were observed in FKBP51-KO animals after acute stress exposure. Following chronic stress exposure, FKBP51-KO mice showed enhanced negative feedback of the HPA axis via lowered baseline Cort levels and more active stress-coping behavior [119]. Moreover, selective silencing of FKBP51 in the amygdala

was found to reduce anxiety-like behavior after acute stress [120], indicating that the influence of FKBP51 on specific behavioral traits is region specific. *Fkbp5* also plays an important role in metabolic challenges, as FKBP51-KO animals show a beneficial outcome when challenged with a high fat diet with an improved insulin tolerance and protection against weight gain [114].

An interesting region to further explore the specific influences of *Fkbp5* is the LC, since this region shows high baseline expression [109], plays a crucial role in initiating the SAM axis, and is connected with the HPA axis via, CRF neurons from several stress-related brain regions that can further innervate the LC [121].

1.5 The Locus Coeruleus

1.5.1 Noradrenergic signaling in the brain: a broad pathway through the neuraxis

The synthesis of NE is restricted to specific brain regions within different parts of the brain stem. The largest, and well-defined brain region containing noradrenergic cells is the LC [122]. The LC was first discovered by Johann Reil in 1809, as he described a "black substance" in the pontine area of the brain, which can be seen without the need of any staining or microscope [123]. The official naming of the LC happened only a few years later when a similar observation by Joseph and Karl Wenzel led to the naming of the nuclei in Latin, "Loci Caerulei", which means "blue spots", and is now referred to in its singular form as "Locus Coeruleus" [124, 125].

The reason for the black coloring is the presence of neuromelanin, which is a dark pigment that is one of the products of the dopamine synthesis, and is therefore present in large quantities in dopaminergic, as well as noradrenergic cells. In the brain this means that neuromelanin is present in the LC, but also in the substantia nigra, which contains a large number of dopaminergic cells. The exact function and role of neuromelanin has been controversial. At first neuromelanin was thought to be a cellular waste product of the dopamine synthesis pathway with no particular function, but nowadays it is linked to specific cellular functions and plays an important role in aging and neurodegenerative diseases, such as Parkinson's and Alzheimer's disease [126, 127].

The LC is a small brain region that is compromised bilaterally of around 3,000 cells in rodents and between 40,000-50,000 cells in humans [128]. The LC is located close to the fourth ventricle in the pontine brainstem [125, 129]. Regardless of its size, the LC is innervated by a wide variety of brain regions with up to 111 afferent brain region connections [130], but even more efferent trajectories, which are globally distributed throughout the entire brain. Distinct afferent trajectories of the LC use different neuromodulators, for instance the prefrontal and anterior cingulate cortices use glutamatergic inputs [131], whereas the posterior lateral hypothalamus uses orexinergic inputs [132]. Importantly, the stress-induced input via CRF-containing pathways is a particularly strong innervation network to the LC, which includes the central nucleus of the amygdala, bed nucleus of stria terminalis, PVN, and several brain stem structures including Barrington's nucleus, and the nucleus paragigantocellularis [133, 134], see Figure 1.7A. In addition, the LC has strong interconnections with other important monoaminergic systems, as it has connections to the serotonin system via the dorsal raphe nucleus [135], and the dopamine system via the ventral tegmental area [136].

The LC has widespread efferent connections throughout the entire neuraxis, including ipsilateral projections to the cortical regions, and bilateral projections to the subcortical and spinal regions. These projection regions include, among others, the frontal cortex, all sensory regions, the thalamic nuclei, and limbic structures, including the amygdala, hippocampus, and septum, but not the basal ganglia [136], see Figure 1.7B. The LC projections have a major effect on a wide variety of different cognitive functions. Early studies utilizing "the loss of function approach" via lesions showed that the LC is involved in the state of wakefulness and arousal [137, 138]. Since then, the LC has been implicated in many other cognitive functions among others, attention shifting and vigilance [139], sensory processing and gating [140, 141], analgesia [142], and learning and memory in the context of fear and aversion [143, 144]. The efferent projections of the LC consist of different subdivided regions within the LC, in which for instance the hippocampus is innervated by the dorsal part, the thalamus by the posterior part, the hypothalamus (including, among others, the PVN) by the anterior part, but the amygdala and cortical regions via a scattered pattern through the entire LC [125, 145, 146]. These efferent NE projections have axons with extensive bifurcations and can travel long distances through many different brain regions, which means that the same NE cell can innervate different brain regions [129, 147]. A long-standing consensus about the LC is that

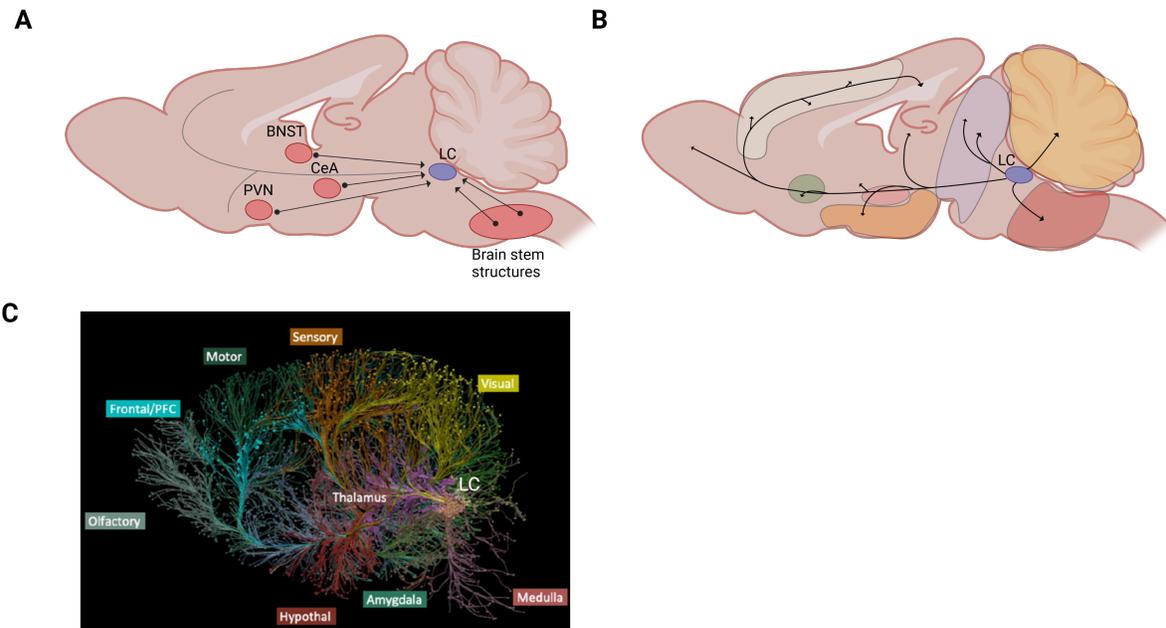


Figure 1.7: **The stress-related input and output pathways of the Locus Coeruleus (LC) throughout the brain.** **A)** The LC is innervated by a wide variety of brain regions with up to 111 afferent brain region connections. In particular, the stress-induced input via CRF-containing pathways is a strong innervation network to the LC, which includes the central nucleus of the amygdala (CeA), bed nucleus of stria terminalis (BNST), paraventricular nucleus of the hypothalamus (PVN), and several brain stem structures including Barrington’s nucleus, and the nucleus paragigantocellularis. **B)** The LC has widespread efferent connections throughout the entire neuraxis, including ipsilateral projections to the cortical regions, and bilateral projections to the subcortical and spinal regions. These projection regions include, among others, the frontal cortex, all sensory regions, the thalamic nuclei, and limbic structures, including the amygdala, hippocampus, and septum, but not the basal ganglia. Due to its widespread connections, the LC has a major effect on a wide variety of different cognitive functions, including wakefulness and arousal, attention shifting and vigilance, sensory processing and gating, analgesia, and learning and memory in the context of fear and aversion. **C)** The axon collaterals of LC neurons are distributed in a coordinated way, per specific cognitive domain. Figure is obtained from [125]

it innervates its projection areas in a uniform way [148]. However, since the last decade, the technological advancement of optogenetic and chemogenetic tools has allowed for the specific manipulation of certain LC projection networks. This has led to novel insights showing a more tailored response of the LC depending on the cognitive domains that are required to be innervated [130, 144, 149]. More specifically, the axon collaterals of LC neurons are distributed in a coordinated way, per specific cognitive domain [125], an example of this distribution was observed by Hirschberg et al. in 2017 [142] that found that the LC-NE projection projecting to the prefrontal cortex that influences aversion and anxiety behavior is innervated separately from the spinal cord projection that influences analgesia, see Figure 1.7C.

Another way the LC can differently influence the cognitive state is via different fundamental modes of activity, which can be either tonic or phasic, which in turn influences the spike pattern discharge and the amount of NE at specific projection sites [150]. The tonic

baseline activity is characterized by a sustained, slow, and regular discharge pattern during awake animals (2–5 Hertz) while being even lower during non-rapid eye movement sleep, and close to completely silent during rapid eye movement sleep [137, 138]. The tonic rates are related to arousal levels within the different states of sleep and awake, in which tonic LC discharge decreases with reduced arousal and disengagement from the environment. Increased tonic discharge in turn increases the state of arousal and goal-directed behaviors [133]. Moreover, high levels of tonic LC discharge are associated with initiating the response towards stressors, which elicits, among others, increased arousal, exploratory behavior, and vigilance [147]. The phasic discharge of the LC is characterized by brief 10–20 Hertz bursts of several action potentials, which are subsequently often followed by a sustained suppression of spontaneous activity (200–500 milliseconds). Phasic bursts are triggered by stimuli that are novel or salient and are not just restricted to stressors, but can also be activated by decision- and response-related signals from prefrontal cortical regions [151].

1.5.2 The role of the Locus Coeruleus in stress-related disorders

The heterogeneity and complexity of the LC are crucial for its role in the stress response system. The LC plays a pivotal role in the processing of acute and chronic stress stimuli via the innervation of the SNS [133, 152], as well as the activation of noradrenergic projections via the entire neuraxis [134, 153–155]. In response to a stressful event, the LC-NE system is activated by several brain regions that use CRF-containing pathways to act predominantly on the CRF_1 receptor within the LC [153]. This activates a state of high-tonic activity in the LC, which in turn activates the LC-NE system and induces a specific behavioral response towards the stressful event [133]. The response of the LC towards stress is broad, and, depending on the type and severity of the stressor, can activate different mechanistic pathways to elicit a distinct and specific behavioral response. More specifically, the central amygdala pathway to the LC can induce a high tonic LC activation that increases brain-wide functional connectivity, which was accompanied by decreased exploratory and increased anxiogenic behavior [134, 156]. The increased state of arousal and cognitive flexibility can be adaptive in life threatening situations [148] but might become maladaptive during prolonged bouts of stress exposure (chronic stress). The sensitivity of LC neurons to CRF is shifted by the exposure to chronic stress, with an increased sensitivity to CRF release, but a decreased

maximum activity response of the LC [157, 158]. The chronic stress-induced sensitization of the LC-NE system can trigger a stress response towards stimuli that would otherwise not have triggered this system. The sensitization of the LC due to stress can contribute to the hypervigilance that characterizes PTSD [133]. Due to the important role of the LC in the brain and the stress response system, it is unsurprising that alterations in the LC-NE system have been linked to stress-related disorders, such as PTSD, but also anxiety-related disorders, and MDD [159–161].

1.6 Rationale and thesis objectives

The lack of a mechanistic neurobiological understanding of stress-related disorders is problematic and remains a global problem, whereas their impact on society is only increasing. Interestingly, exposure to chronic stress does not lead to the development of psychopathology in all individuals, which is due to the combined effect of genetics, epigenetics, and the environment. The psychiatric risk factor FKBP51 has been shown to be involved in adapting the vulnerability towards stress-related disorders in clinical as well as preclinical research, but the region-specific effects on the behavioral symptoms remain elusive. The major role of the LC in both stress response systems, together with its wide projections throughout the neuraxis, and the relatively high expression of FKBP51, make the LC a promising region to further explore and elaborate the role of FKBP51 in the stress response system. However, the preclinical tools to assess the behavioral symptoms of stress-related disorders such as social aversion, have become controversial. They rely on a reductionistic approach in which animals are tested in an instrumental environment, that often does not reflect the ethologically relevant behavioral constructs. This has limited the utility of behavioral experiments and could lead to false interpretations of the data. Therefore, the current thesis aims to provide a novel behavioral assessment tool for preclinical social behavior research using state-of-the-art computational tools to explore the role of stress on behavior and the manipulation of FKBP51 in the LC.

In chapter 2 the history of the social behavioral assessment and how machine learning tools can advance the field of social behavior are reviewed. In chapter 3 a novel tool to assess social behavior, called "DeepOF" is introduced and the different possibilities are investigated for the social behavioral classification, using both supervised and unsupervised classification

tools. This newly developed deep phenotyping tool is then further utilized in chapter 4 to explore the social behavioral profile after exposure to chronic social defeat stress in male mice, and in chapter 5 to explore a more in-depth analysis of the sex-dependent early life stress effects on fear conditioning. Further focusing on the genetic risk factor, chapter 6 combines DeepOF behavioral analysis with molecular profiling to explore the role of FKBP51 in the LC on stress exposure and social behavior. Finally, in chapter 7 a perspective on the stress-related research field is highlighted by specifically investigating the integration of novel advanced molecular and computational tools.

Chapter 2

Advancing social behavioral neuroscience by integrating ethology and comparative psychology methods through machine learning

Bordes, J., Miranda, L., Müller-Myhsok, B., Schmidt, M.V.

Originally published in:

Neuroscience and Biobehavioral Reviews, 2023, 151, 105243



Advancing social behavioral neuroscience by integrating ethology and comparative psychology methods through machine learning

Joeri Bordes^a, Lucas Miranda^{b,c}, Bertram Müller-Myhsok^b, Mathias V. Schmidt^{a,*}

^a Research Group Neurobiology of Stress Resilience, Max Planck Institute of Psychiatry, 80804 Munich, Germany

^b Research Group Statistical Genetics, Max Planck Institute of Psychiatry, 80804 Munich, Germany

^c International Max Planck Research School for Translational Psychiatry (IMPRS-TP), 80804 Munich, Germany

ARTICLE INFO

Keywords:

Social behavior
Psychiatric disorders
Ethology
Comparative psychology
Stress

ABSTRACT

Social behavior is naturally occurring in vertebrate species, which holds a strong evolutionary component and is crucial for the normal development and survival of individuals throughout life. Behavioral neuroscience has seen different influential methods for social behavioral phenotyping. The ethological research approach has extensively investigated social behavior in natural habitats, while the comparative psychology approach was developed utilizing standardized and univariate social behavioral tests. The development of advanced and precise tracking tools, together with post-tracking analysis packages, has recently enabled a novel behavioral phenotyping method, that includes the strengths of both approaches. The implementation of such methods will be beneficial for fundamental social behavioral research but will also enable an increased understanding of the influences of many different factors that can influence social behavior, such as stress exposure. Furthermore, future research will increase the number of data modalities, such as sensory, physiological, and neuronal activity data, and will thereby significantly enhance our understanding of the biological basis of social behavior and guide intervention strategies for behavioral abnormalities in psychiatric disorders.

1. Introduction

Understanding behavior in humans and animals is essential, as it forms a central aspect of our existence and interaction with the environment (Heimlich and Ardoin, 2008; Bolhuis et al., 2021). At the same time, behavioral abnormalities are at the core of severe and highly prevalent psychiatric disorders. Specifically, the impairment of social functioning is an important symptom of many different psychiatric disorders, such as post-traumatic stress disorder (PTSD), depression, autism, and schizophrenia (Peleh et al., 2019a; American Psychiatric Association, 2013; Nietlisbach and Maercker, 2009; Dodell-Feder et al., 2015; Katz et al., 2011). Therefore, the accurate measurement of different social behaviors is crucial for a better understanding of psychiatric patients and contributes to solving the complex mechanisms of those disorders. Preclinical rodent models are widely used to investigate the underlying mechanisms of complex social behaviors (Peleh et al., 2019a; Beery and Kaufer, 2015), as well as to evaluate the effectiveness of novel drug interventions (Pöhlmann et al., 2018; Lopez et al., 2022). However, the assessment of social behavior as a holistic construct is challenging as it relies on many different types of social behaviors and is

influenced by different factors, such as stress exposure (Beery and Kaufer, 2015). Therefore, a variety of behavioral tasks have been employed to understand different mechanisms of the social behavioral construct in rodents (Toth and Neumann, 2013; Fan et al., 2019; Zhou et al., 1979; Golden et al., 2011).

1.1. The history of animal behavior

Exposure to external stimuli triggers a distinct pattern of cellular responses in the brain, which ultimately drives the behavioral response. Fundamental concepts of behavior have historically been established using observational research, where animals are left undisturbed and observed in their natural habitat. One of the earliest and most influential animal behavioral research was performed by Charles Darwin in the 1800s (Thierry, 2010), where he observed variations between animals and concluded that species descended from each other (Darwin, 1859). This type of behavioral research led to the development of the field of ethology, which concentrated on the evolutionary basis of animal behavior while using predominantly observational approaches. An overarching principle in ethology is that behavior is described in its

* Correspondence to: Research Group Neurobiology of Stress Resilience, Max Planck Institute of Psychiatry, Kraepelinstr. 2-10, 80804 Munich, Germany.
E-mail address: mschmidt@psych.mpg.de (M.V. Schmidt).

highest complexity by observation in natural or semi-natural environments. This allows the researcher to study behavior in a descriptive way, using a hypothesis-generating approach to discover novel behavioral concepts. Important examples of outstanding work in ethology include the research of Tinbergen and colleagues, who observed that young birds can distinguish the shape of a hawk and a goose and change their behavior according to the perceived shape (Tinbergen, 1939). In addition, the work of Lorenz and colleagues observed an interesting phenomenon in young domestic chicks, where young animals narrow their social preference towards an object during early exposure, which typically is their parent (Bolhuis and Giraldeau, 2005; Sulloway, 1982). The field of ethology and observational research has led to phenomenal research findings, but there are also important limitations in the study design. Observational studies rely on the researcher's ability to assess behavior, which can lead to misinterpretation and consequently differences between researchers in the interpretation of the observed behaviors. In addition, the lack of control over the environment limits the throughput and the research question that can be assessed, which can lead to unreproducible results due to high variability between the experimental conditions.

In response to these limitations, the field of comparative psychology was introduced, in which researchers aimed to study and understand behavior using tightly controlled environmental settings. Here, the focus was on breaking the behavioral construct down into clearly identifiable and quantifiable syllables by reducing the complexity of the behavior and limiting the dimensions that behavior can be expressed in through precise laboratory settings. These laboratory tasks are characterized by a high controllability over the environment and a standardized set of behavioral read-outs and therefore require hypothesis-driven research questions. Several outstanding researchers showed the novelty and strength of this field, including Thorndike, Skinner, and many others. Thorndike demonstrated the construct of trial-and-error learning, as he observed that by increasing the number of trials, animals became more efficient to escape an apparatus and gain their reward (Thorndike, 1898). Subsequently, Skinner developed one of the first and extremely popular behavioral laboratory tasks, the operant conditioning chamber (Skinner, 1948). The task can be used for negative and positive reinforcement learning and is still used in research around the globe to this day forward, underlining its excellence. By this time, a trend to standardize and simplify different behavioral disciplines using laboratory tasks was set in motion (Tolman, 1948; Aggleton, 1985; Barnes, 1979; Morris, 1981), and is currently still ongoing. The excellence of laboratory tasks for behavioral science is undoubtedly great, especially for investigating the influence of external stimuli (e.g., the stress response system) and the role of interventions (e.g., pharmacological) on the behavioral output (Hånell and Marklund, 2014). In addition, the unprecedented options for using different genetic mouse models have enabled the exploration of specific target genes on behavior. Unfortunately, no behavioral tasks are without flaws, and there are specific concerns regarding laboratory tasks that need to be addressed. The laboratory set-up provides an intensive interaction of the researcher with the testing animals, which causes concerns, as inter-individual differences between researchers can influence the outcome of the behavioral performance of the animals (Sorge et al., 2014; Chesler et al., 2002). As an example, a recent study described the influence between male and female experimenters on the largely different behavioral outcomes of drug-treatment efficacy (Georgiou et al., 2022). Another concern is the difficulty of streamlining animal housing and testing. Even though compared to observational research the differences in housing and testing are subtle, they have been found to lead to different behavioral outcomes (Richter et al., 2009), and therefore can give a false sense of reliability. In addition, although group housing of laboratory mice is crucial for their social welfare, it remains a highly unnatural environmental setting and has its limitations. The confined space with several same-sex mice hampers the animals in expressing certain species-typical behaviors, such as the search for food and water, the

vigilance for certain threats, and the motivation for sexual reproduction. In addition, the lack of escaping introduces changes in territorial aggression, as, after an aggressive encounter, the dominant and subordinate animals cannot go their separate paths and will have to stay confronted with each other, which introduces problematic behavior, such as severe aggression (Weber et al., 2017). Finally, the use of inbred mice is a problematic development for investigating naturalistic behaviors. While the genetic models have allowed for important insights into the genome, they have left us with an animal model that behaves rather differently than their conspecifics in the wild, which questions the validity of the animal model and unfortunately has reduced the reproducibility of the behavioral research (Richter, 2017; Wahlsten et al., 2003; Voelkl et al., 2020; Crawley, 1996).

The reductionistic approach in laboratory tasks is a strength for many different behavioral disciplines; however, it can be a pitfall for behavioral constructs that rely on many different behavioral outputs and require more naturalistic environments, which therefore are more complex to assess (Blumstein, 2010). In such cases, laboratory tasks can oversimplify behavior, which can lead to an inadequate or even wrong assessment, that is often lowering the reproducibility rate. A good example of such a complex behavioral construct is social behavior, therefore the next section describes the definition of social behavior, together with the current behavioral tasks that are used to measure social behavior.

1.2. The social behavioral construct

Social behavior is a naturally occurring construct in many different species, including rodents and humans, and holds a strong evolutionary component, as it is critical for the survival of the individual as well as for the species (Wei et al., 2021; Chen and Hong, 2018). These behaviors are innate but are strongly dependent on the environmental circumstances and will adapt the social behavioral output appropriately. For example, adolescent mice will indulge in social play interactions with their conspecifics, but if they are socially isolated during this time, their social behavioral repertoire and the underlying neurobiological mechanisms will be altered during adult age (Musardo et al., 2022). Social behavior is a complex construct and is used as an umbrella term for a plethora of different behaviors, which makes its precise definition challenging. In general, social behavior has been defined as all behaviors that influence or are influenced by other members of the same species. This terminology includes sexual and reproductive activities, all behaviors that bring individuals together, as well as aggressive and dominant behaviors (Chen and Hong, 2018; Whishaw et al., 2006; Mackintosh and Grant, 1963). Although not the focus of this review, it is important to note that this definition excludes social interactions across different species, or objects, which are part of the social behavioral repertoire, even though they might be the minority of the social encounters.

2. Social behavioral tasks

Social behavioral events in rodents typically begin with the detection of a social encounter. The animal can decide to ignore or avoid the social encounter, or to consummate the social behavioral encounter, which relies on one or a combination of different social behaviors, such as parenting, fighting, sexual interaction, following, feeding, playing, and guarding (Wei et al., 2021). These constructs are influenced by other neuronal modalities, including motivation (Bariselli et al., 2018; Solié et al., 2022a), emotion (Sakaguchi et al., 2018; Pisansky et al., 2017; Keum and Shin, 2019; Jabarin et al., 2022), and decision-making (Scheggia et al., 2022), which means that the underlying social behavioral construct is complex to assess, as it relies on many different behavioral read-outs. Currently, the assessment of social behavior relies on using a variety of behavioral tasks measuring different social domains, see Box 1. The majority of the tasks have been developed with the goal of increasing the throughput and the environmental controllability,

Box 1

The different social behavioral domains and their behavioral tasks.

The assessment of social behavior relies on a variety of behavioral tasks that can measure different social domains, including dominance, avoidance, approach memory, transmission, decision-making, and conditioning. Dominance behavior is generally assessed using competition tasks, such as the tube test (Fan et al., 2019; Lindzey et al., 1966) and the warm spot test (Zhou et al., 1979), or via long-term home cage observation (Shemesh et al., 2013; Forkosh et al., 2019). The avoidance and approach behavior can be assessed using tasks that compare the interaction with a non-social stimulus and a social stimulus, using the three-chamber task (Rein et al., 2020) and the social avoidance task (Berton et al., 2006; Lukas et al., 2011), but also with freely-moving multi-animal social interaction tasks (Bordes et al., 2022). In addition, social memory can also be assessed in the three-chamber task and social avoidance task, but then compares a familiar-social stimulus with a novel social stimulus (Gheusi et al., 1994; Winslow, 2003; Engelmann et al., 2011). The transmission of social behavior is a particularly complex domain, which can be subdivided into many different subfields. One example is the social transmission of food preference, in which a naïve mouse follows a demonstrator mouse who has previous knowledge of a food reward (Wrenn, 2004). In the context of stress, a prominent domain is social fear transmission, in which animals receive shocks when interacting with a social conspecific (Toth et al., 2013), or observe other animals getting foot shocks (Jeon and Shin, 2011). The fields of decision-making and conditioning have recently been further adapted to the social behavioral research field and interestingly behavioral tasks have been adapted to specifically measure these domains in a social context, using the two-choice social decision-making task (Scheggia et al., 2022), and the social instrumental task (Solié et al., 2022b).

but therefore rely on a reductionistic approach using single or few behavioral read-outs (univariate behavioral tests), a hallmark of the comparative psychology field. The next section describes the social behavioral domains of social dominance and social avoidance. Although many more variants of social behavior and corresponding tasks exist, these social behavioral domains in particular are of interest, as they are well-established behavioral tasks in the field of stress neuroscience and they showcase the development of the reductionistic approach in social behavioral tasks.

3. Social dominance

An important domain of social behavior is the social hierarchy, which is important for a balanced and structured group of multiple animals from the same species (Dwartz et al., 2022; Qu et al., 2017). Social status impacts the chances of survival and, among other aspects, the individual's reproductive chances, health, and food availability (Tamashiro et al., 2005; Ellis, 1995; Sapolsky, 2005). The social status is conveyed via dominant, aggressive behaviors, and the subsequent submissive, defensive behaviors, which are adaptive behaviors to reduce aggression in established hierarchies. The dominant animal is commonly identified by consistent wins in social conflicts with the subordinate animals (Fan et al., 2019; Dwartz et al., 2022; Wang et al., 2014). Several behavioral tasks have been developed to investigate the hierarchical structure in rodents and the most popular tests include the tube test (Fan et al., 2019; Lindzey et al., 1966) and the warm spot test (Zhou et al., 1979), but also via long-term home cage observations (Shemesh et al., 2013; Forkosh et al., 2019). Other hierarchy tests rely on the competition for food or water, but those tests have decreased in popularity, as they are influenced by the motivation for food and water and can fluctuate based on saturation (Merlot et al., 2004; Cordero and Sandi, 2007; Ujita et al., 2018). The tube test has gained popularity, due to its simplicity and an available standardized protocol by Fan et al., (Fan et al., 2019). The tube test utilizes a narrow transparent tube, in which mice are placed on opposite ends. Upon release, the mice need to go through the tube towards the other side, but can only succeed by forcing the conspecific animal to go backward. The subordinate mouse is identified as the animal that has to go backward, whereas the dominant mouse is the animal that successfully pushes the conspecific away. The tube test relies on the assumption that after several training days, the animals prefer to go to the other side of the tube and use dominant-related mechanisms to establish, which of the animals back out. The animals are tested over several consecutive days, after which a stable and consistent hierarchy can be observed (Fan et al., 2019). However, since the animals are confined within the tube, their agility will influence their winning chances (Fan et al., 2019; Zhou et al., 2018),

which can be different due to many factors, such as age, strain, stress levels, weight, etc. Moreover, the tube test can only test two mice at the same time, which limits the transferability towards a more naturalistic environment, where animals interact and are challenged within group dynamics. A recently developed hierarchical test in mice that overcame this problem is the warm spot test, which works reliably with up to four mice. In this task, the animals are placed on a cold floor with a warm spot in the corner, which is only big enough to warm one animal at a time (Zhou et al., 1979). The test relies on the behavioral construct that mice prefer the warm spot over the cold area and will determine the time spent at the warm spot based on their rank in the hierarchy. The hierarchy state of animals is observed during a 20-minute test, in which the amount of time spent in the warm spot should overlap with the hierarchical position of the animals. Therefore, the dominant animal spends the most time in the warm spot, whereas with every lower ranking, each animal spends less time in the warm spot (Zhou et al., 1979). Most importantly, the tube test and warm spot test show a similar pattern of hierarchical structure (Zhou et al., 1979) and are highly correlating with other markers for dominance, including, among others, agonistic-related behaviors, urine markings, and ultrasonic vocalizations (Zhou et al., 2018; Wang et al., 2011). These tests have contributed significantly to our understanding of hierarchy in mice, are well appreciated for their simplicity and robustness, and have been widely adopted for stress-related research (LeClair et al., 2021; Park et al., 2018; Larrieu et al., 2017; Matthews et al., 2016). However, these tests rely on an instrumental laboratory environment that is particularly unnaturalistic. In the tube test animals are forced within a confined space towards confrontation, and in the warm spot test, the animals are forced into confrontation as there is only one place that is pleasantly warm enough. Therefore, these tests exclude the natural influences of motivation on hierarchy establishment. Several studies have shown individual differences in the motivation to pursue social interaction with conspecifics (Bariselli et al., 2018; Torquet et al., 2018), which includes aggressive confrontation between animals. Thus, these tests reflect the reductionist approach, as they simplify the behavioral read-out.

4. Social avoidance behavior

Clinical research has widely recognized that social withdrawal and avoidance are key symptoms of depression (Ottenbreit et al., 2014; Fernández-Theoduloz et al., 2019) and anxiety disorders (Rinck et al., 2010; Heuer et al., 2007), which is why understanding the underlying mechanisms related to social avoidance is crucial. The domain of social behavior containing social avoidance is an often-used paradigm for investigating the stress response system and in particular, has been well-established to assess stress resiliency versus susceptibility of

individual rodents (Krishnan et al., 2007; Murra et al., 2022; Donahue et al., 2014), but is also used in many other scientific fields to investigate the affectedness of social behavior, for instance in autism (Rein et al., 2021; Rapanelli et al., 2021; Bidinosti et al., 2016; Qin et al., 2018). Therefore, several behavioral tasks have been developed to measure social avoidance behavior (Toth and Neumann, 2013), for example, well-established tasks include the three-chamber task (Rein et al., 2020) and the social avoidance task (Berton et al., 2006; Lukas et al., 2011). These tests have been described in detail in other studies (Rein et al., 2020; Berton et al., 2006; Lukas et al., 2011; Zimprich et al., 2017). Both tests rely on the exploratory behavior of the test mouse towards a novel social stimulus. The social stimulus is an unknown social conspecific, who is constrained in a small grid box or beaker and the interaction time of the test subject is measured in relation to the interaction of an empty box directly before the social interaction test (social avoidance test) or at the same time in a separate chamber of the apparatus (three-chamber test). The gold standard for the calculation of social avoidance behavior relies on the interaction ratio (IR) (Krishnan et al., 2007; Donahue et al., 2014). The IR is calculated by taking the total amount of time spent with the social stimulus (T_S) and dividing it by the total amount of time spent with the empty cage (T_E) resulting in the following formula: $IR = T_S / T_E$ (this score can also be expressed in a percentage which would mean the end result is multiplied by "100"). The expectation is that unaffected social behavior in mice results in an IR higher than "1", which means that animals spend more time investigating the social stimulus compared to the empty cage. In the same line of thought, animals that have an IR lower than "1" show an avoidance of the social stimulus and therefore have affected social behavior. These tasks allow for a quick and simple analysis of social avoidance behavior and enable high controllability over the environment. In the last decade, they have gained popularity and have been extensively used for observing social avoidance-related behaviors specifically in stress research, where they have shown their value (Murra et al., 2022; Reguilón et al., 2021; Durand-de Cuttoli et al., 2022; Morel et al., 2022; Li et al., 2022). However, these tests also rely on a reductionistic and instrumental environment, as the social stimulus is trapped in a confined space, which increases the risk for cross-over effects by other types of behaviors, such as motivation, emotion, and anxiety. Moreover, the robustness of these tasks has been questioned, due to a lack of standardized experimental set-ups and subsequently high variation in the behavior of individual animals (Pearson et al., 2010). A critical problem that has been overlooked until now is the fact that there is a proportion of nonstressed animals that are also showing social avoidance behavior towards the social stimulus (e.g., 12% (Krishnan et al., 2007), 26% (Golden et al., 2011)). This means that there is a subset of stressed animals that are categorized as susceptible ($IR < 1$), even though they showed avoidance behavior already before the stress exposure. Therefore, the assumption that nonstressed animals prefer the novel social stimulus over non-social environments seems to be highly dependent on the individual. Even though these animals are an extremely interesting group to investigate their social behavioral profile, they cannot be categorized in the same group as animals that initially do show social preference but change upon stress exposure. This can be avoided by screening all animals before the stress exposure so that animals that are already avoiding the social stimulus can be investigated as a separate group or excluded.

Measuring constructs of social behavior using a reductionistic approach has its advantages, as it allows for high controllability over the environment, straightforward comparability across cohorts and conditions, high throughput, and simple analysis of the behavioral read-outs. These reductionistic social behavioral tests have provided important insights, especially into the stress-related mechanisms of social behavior, but there are severe limitations that need to be addressed. These tests often suffer from problems related to the comparative psychology, reductionistic research approach, which means that some of these tasks have limited ecological validity, as they largely restrict free social interaction behavior (such as fighting, following, grooming,

anogenital sniffing, etc.) and only measure one or a few simple behavioral read-outs (e.g. the preference of an animal to be in close proximity of a conspecific). Therefore, these tasks tend to oversimplify the social behavioral repertoire and do not accurately capture the full range of the social behavioral construct. In addition, these tasks have been developed specifically for rodents and have limited comparison to human behavioral tasks, which lowers the translatability to the clinic. Therefore, the development of more naturalistic types of behavioral tasks, including free social interaction between multiple animals, will be crucial for a more accurate and ethologically relevant assessment of social behavior. However, up until recently, no software tools were available to analyze and distinguish multi-animal interactions, but the many different behaviors within a free social interaction task are too complicated, time-intensive, and repetitive to assess manually (Hånell and Marklund, 2014; Goodwin et al., 2020; Sturman et al., 2020). The rapid advances in automatically annotated motion tracking opened the possibility to assess many different social behaviors in complex environments while maintaining high throughput. Therefore, the next section will discuss the current status of motion tracking and machine learning for behavioral classifications, and their implications for social behavioral neuroscience research.

4.1. Automatically annotated motion tracking and machine learning tools advance the behavioral analysis

When looking back at the history of the development of behavioral assessment, there were several influential methods using both observational research approaches in natural environments as well as the reductionistic laboratory research approach. Both methods have their inherent advantages and disadvantages, but by combining the strengths of both methods, it is possible to overcome some of the current problems with (social) behavioral assessment (Fig. 1). This can be achieved by using a semi-naturalistic environment, in which animals can freely interact with each other, but are still limited in their environment and space, to maintain a certain level of control over the environment. Several researchers have contributed to the development of such tasks (Shemesh et al., 2013; de Chaumont et al., 2019), which contain automatic phenotyping systems using top-view camera detection systems, in which up to four mice can live and freely interact for many days without experimenter intervention. Shemesh et al., use different fur color dyes to distinguish the animals, which therefore allows video recordings only with white furred animals (e.g. CD1, BALB/c), as with darker fur (e.g. C57/Bl6) the tracking system cannot distinguish well enough (Shemesh et al., 2013). This is an important drawback as many studies are conducted with C57/Bl6 animals. This problem was solved by combining video recordings with RFID probe identification of the animals via detection antennas under the arena, which therefore does not require different colorings for the animals (de Chaumont et al., 2019; Peleh et al., 2019b). However, the RFID probes are expensive, cannot be reused, and have problems when animals are in close proximity together, making the technique difficult to implement in many different labs. In addition, for both methods, the animals need to be sedated to color dye or implant the RFID probe, which can be seen as a stressor and is therefore not an optimal solution for investigating, e.g. consequences of stress exposure. The tracking and analysis in these systems remain a difficult process and rely on homemade tailored software that is likely too complicated to integrate into many different labs. Also, the tracking of the animals consists of only the center-point of the animals and at best includes the tail base and nose. This limits the information and accuracy that can be obtained about the animal's behavioral profile, especially looking at more complex social behaviors (such as chasing, mounting, and attacks), but also individualistic behaviors (such as grooming, huddling, digging, and sniffing).

In 2018, Mathis et al., launched the open-source software system, DeepLabCut (DLC) (Mathis et al., 2018), which allowed easy access for researchers to track animals with high throughput for any given

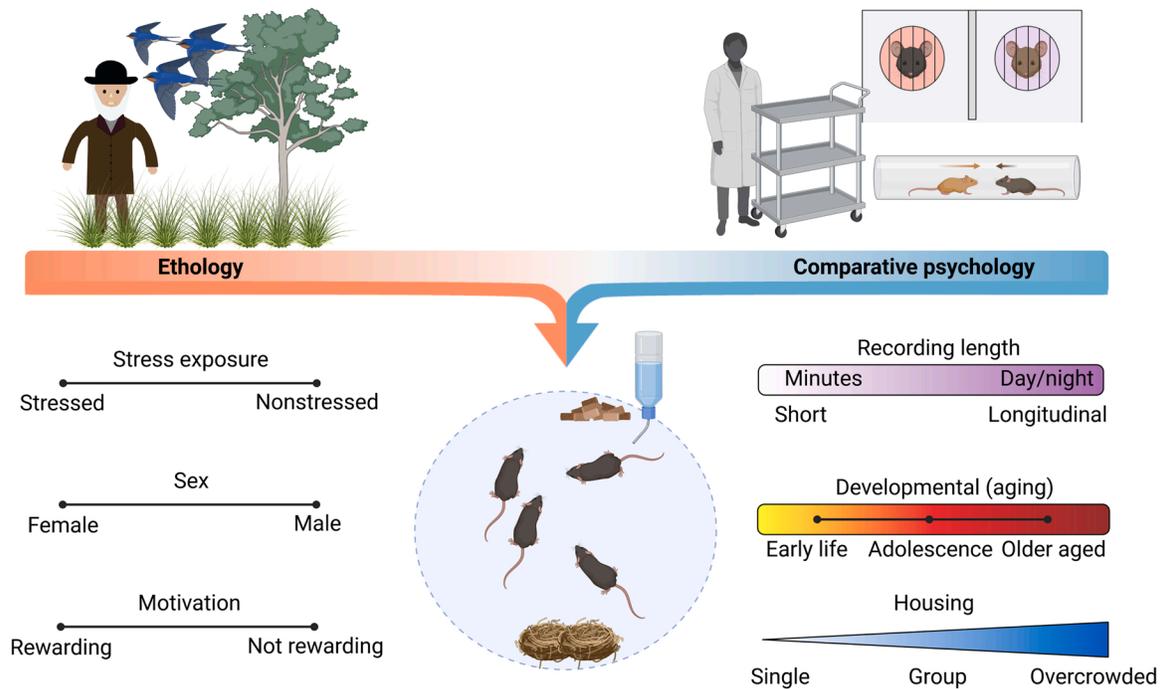


Fig. 1. Next-generation of social behavioral tests. In the top part of the figure the important methods for investigating behavioral neuroscience are illustrated, with ethology on the left and comparative psychology on the right. Next-generation social behavioral tests (illustrated by the arrow in the middle) rely on a semi-naturalistic environment, which includes the strengths of both ethology, in which a naturalistic set-up without the interference of the experimenter is implemented, as well as comparative psychology, whereas certain controllability of the environment remains by limiting the space and influences from the outside world. Different factors, such as stress exposure, sex, motivation, recording length, developmental (aging), and housing conditions, will influence the outcome of the social behavioral assessment.

behavioral video recording. DLC utilizes deep neural networks that enable quick and easy markerless pose estimation and only need little training data to achieve human-level accuracy for object recognition (Mathis et al., 2018). Up until then, the gold standard for tracking animals relied on commercial software systems that could only track the center point of single animal video recordings. The analysis of social interaction data between multiple animals remained to be assessed manually, but as social behavior relies on many different behavioral interactions, this limited the complexity of the behaviors that could be assessed (Sturman et al., 2020; Bordes et al., 2022). DLC enables the tracking of any body part of interest across multiple animals, without the need for additional invasive marking methods (such as RFID tags or animal dye coloring), as the neural networks can also separate identically looking animals (Lauer et al., 2022). Importantly, to gain accurate multi-animal tracking, the DLC models for tracking multiple animals load entire video fragments onto the graphics processing unit (GPU) during training, as opposed to single frames in the single-animal models, which requires substantially more advanced hardware. Since then, other open-source markerless pose estimation software systems have become available, such as DeepPoseKit (Graving et al., 2019), Social Estimates Animal Poses (SLEAP) (Pereira et al., 2019, 2022), and SIPEC (Marks et al., 2022). These different software systems acquire pose estimation in slightly different ways and therefore vary in their accuracy, training speed, and amount of required training data (Mathis and Mathis, 2020). Ultimately, however, they all are open-source tools that allow fast and easy tracking of any body part of interest in single or multiple animal recordings. Their offered features and performance are to date comparable, and choice often relies on community support and downstream analysis compatibility. Importantly, DLC has an extensive user base and has shown a strong dedication to supporting, maintaining, and extending its software ecosystem with new downstream analysis tools (Mathis, 2020; Schneider et al., 2023). These pose-estimation tools have enabled the social behavioral research field to be able to investigate more

naturalistic types of behaviors, as the tracking of many freely moving animals can easily be performed. However, this provides a novel problem, as pose-estimation tools increase the amount of data tremendously. The interpretation of the pose-estimation output, the “X and Y” coordinates over time, needs to be translated into meaningful behaviors. The most basic analyses include tracking the center point of the animals as described before, but in a markerless way, which contributes to automating the quantification of previously used univariate tests (Sturman et al., 2020). However, tracking multiple body parts enables researchers to explore a plethora of different behaviors based on a more complete capture of posture. Different analysis methods can be employed to analyze the pose-estimation data using supervised and unsupervised analysis toolboxes.

The supervised analysis toolbox can be employed using machine learning models or simple rules to identify specific posture-based behaviors of interest. An important package that uses a minimal amount of coding to create supervised behaviors from DLC and other pose-estimation packages, is SimBA (Simple Behavioral Analysis (Nilsson et al., 2020)). SimBA allows practitioners to label behaviors of interest in a graphical user interface to train machine learning models that can learn the rules governing these patterns from data, automating the quantification of arbitrarily complex traits. The provided models are based on extracted static and dynamic features describing animal motion, instead of on the sequences themselves. This makes the models easier to train and does not require a dedicated GPU, but the transferability toward other behavioral setups is rather limited. However, some behaviors do not need machine learning models as they can accurately be deduced via hardcoded rules (Bordes et al., 2022). These can include, but are not limited to, time-in-zone quantifications, certain interactions between individuals (for example, nose to nose and nose to tail), or interactions with objects. Moreover, certain packages, such as SIPEC (Marks et al., 2022) or MARS (Segalin et al., 2021), combine a tracking system with a pipeline for supervised behavioral annotation,

which can show advantages over combining different packages to the same end, since users do not need to worry about software compatibility as everything is contained within the same framework. The current machine learning models that are used in many of the different supervised analysis toolboxes, tend to overfit the nuances present in the dataset they were trained on, which limits the transferability to a new data set. Transferability is an important way to standardize the annotated behavior of interest and allows a minimal amount of work for other labs to implement the tagging of previously annotated behaviors. An important future goal will therefore be to find models that can extend their detection with greater accuracy to completely novel data sets.

Another method to examine behavior is through studying pose-estimation data without prior categorization. This can be performed using unsupervised learning, which is a branch of machine learning that aims to extract insights from the data without using information about the behaviors of interest a priori. This method segments animal trajectories across time to recover consistent behavioral patterns, which can coincide with the patterns that were analyzed in a supervised manner. In addition, this method allows for a hypothesis-generating approach, as it can identify behavioral patterns that indicate novel behaviors in the specific behavioral context. Since unsupervised approaches allow to explore the behavioral space without the need for time-intensive labeling, they can be used as an initial screening for behaviors of interest. For example, unsupervised analysis can be used to identify specific behavioral patterns that are showing the biggest difference between a set of defined experimental conditions. As a next step, researchers can then train supervised classifiers to measure the behavior of interest more directly and with less noise, or to initialize classifiers that are further

fine-tuned with human feedback, using an active learning approach (Yttri et al., 2023). Several packages and pipelines have come out that use clever ways to segment behavior in an unsupervised way. The software system B-SOiD (Hsu and Yttri, 2021), for example, relies on annotating the motion data with sets of features that can help describe behavior across time without using sequential data directly. Another software system, MoSeq (Wiltschko et al., 2020), takes advantage of the time component of motion using autoregressive hidden Markov models, which can directly capture probabilistic relationships between the input variables. While originally devised for depth sensing camera setups, recent iterations adapt it to pose estimation settings too (Weinreb et al., 2023). Other packages, such as VAME (Luxem et al., 2022) and DeepOF (Bordes et al., 2022) use neural networks that can process the motion sequences directly, which was shown to provide more meaningful clusters of specific behaviors with less noise of other behavioral patterns (Luxem et al., 2022). Furthermore, an advantage of neural network models is that, in parallel to sequence segmentation, they allow researchers to embed motion trajectories into interpretable latent spaces that can be analyzed, for example, differentially across experimental conditions. Moreover, the explainability of the retrieved patterns is key to understanding the underlying behaviors (Goodwin et al., 2022). Being able to interpret how and what the models are doing can help improve transferability across datasets on the one hand, but also categorize unsupervised clusters on the other hand, and this can be done both via visual exploration (taking advantage of video data and going back to the snippets that are assigned to specific patterns) or using machine learning explainability tools, such as Shapley Additive Explanations (SHAP) (Goodwin et al., 2022). A comprehensive overview of

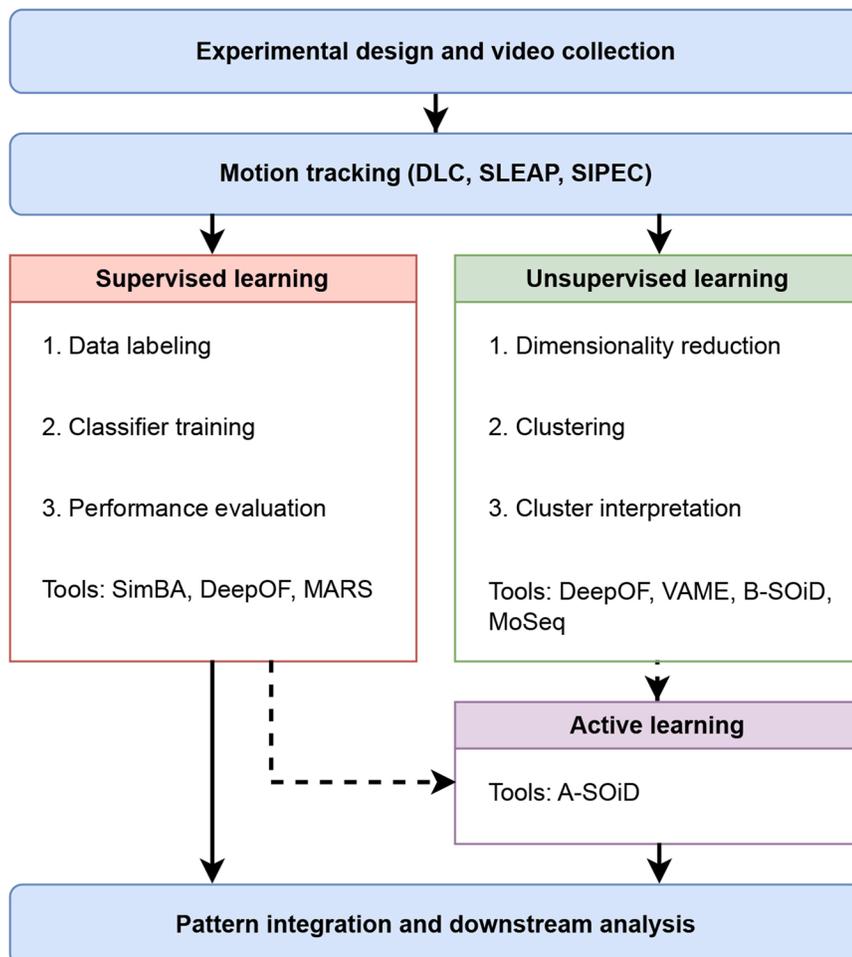


Fig. 2. Automatic behavioral annotation via motion tracking. After the experimental design is defined and videos are collected, key points over time for one or more animals are extracted as time series, using pose-estimation tools, such as DLC, SLEAP, or SIPEC. Predefined behaviors can then be extracted using supervised learning tools such as SimBA or MARS, which typically require data labeling, followed by classifier training and performance evaluation of the extracted behaviors. Other tools, such as DeepOF, provide pre-trained models that bypass these steps. Another way of extracting broader information is unsupervised learning, which does not require labeling, and aims to obtain behavioral syllables (or clusters) after a dimensionality reduction step. Cluster interpretation is a key step, which can be achieved by visual video inspection or using model explainability tools such as SHAP. Tools in this realm include DeepOF, VAME, B-SOiD, and MoSeq. Moreover, unsupervised learning results can be used to initialize supervised models with active human feedback, such as seen on the A-SOiD framework. Finally, the expression and dynamics of all retrieved patterns can be compared across experimental conditions, for example, to gain insights on behavioral shifts.

this pipeline is presented in Fig. 2.

In contrast to the aforementioned open-source software packages, it is worth mentioning that there are also commercial solutions that have been employed for animal behavioral analysis, such as Smartcube (Lorsch et al., 2021). The promise of delivering an out-of-the-box, easy and high-quality product for a price is valid, however, there are intrinsic disadvantages to using proprietary black-box solutions. Commercial software systems are, for understandable reasons, not subjected to the same rules as scientific software, which means that they are not peer-reviewed and the information about which and how the employed models are working is not openly available. A potential way for bypassing those problems is to visually validate the produced results, which may in turn require more manual labeling. In the end, the fast solution-based development that the open-source community has shown in recent years is hard to compete against. An example of this fast development is the several publicly funded open-source competitions that were released in recent years, with the aim of improving machine-learning algorithms for behavior (Sun et al., 2021). This has the potential to continuously improve the current state of the art and yield a constant stream of open-source tools of excellent quality for behavioral research. Moreover, all data produced by labs while using this type of software package are collected and recycled to further improve previous algorithms, creating a constructive feedback loop from which the whole community can benefit.

Although the field of automatic animal tracking and behavioral annotation has progressed unprecedentedly in the last few years, there are still several issues to overcome (von Ziegler et al., 2021). The increased number of open-source packages to analyze behavioral data is a crucial development that creates novel and better models, but many packages offer similar and overlapping results that are relying on different software systems and are therefore often not compatible with each other. This problem limits users to find the best set of software systems that are needed to answer their research question, especially for researchers from non-computational labs. An important advancement by the DLC development team is the future implementation of pre-trained tracking models that do not require labeling of any data and can export a flexible set of annotated body parts, which means that if competing software packages require a different set of labeled body parts, this would easily be extracted using the DLC Model Zoo Super-Animal software system (Ye et al., 2022). The current fast trajectory of many newly developed software solutions for behavioral analysis is encouraging, but in order for a wide audience to implement these tools, it would be beneficial if the field will converge to a centralized solution that ensures a compatible ecosystem for tracking, annotation, and the embedding of multimodal behavioral data, to which researchers would be able to openly contribute.

The development of more precise tracking tools together with the numerous post-tracking analysis packages have allowed the behavioral research field, and in particular social behavioral research, to rethink the possibilities for the assessment of behavior. The implementation of such tasks will be beneficial for many different neuroscientific behavioral fields including fundamental behavioral research, especially studies focused on the consequences of stress exposure. An increasing number of studies are using these tools to investigate how stress affects social behavior. A selection of those studies is highlighted in the next section to illustrate the possibilities and strengths of such tasks using novel deep-learning tracking tools.

4.2. Investigating social behavior using advanced tracking tools for better stress phenotyping

Recently, several studies have used advanced tracking tools to investigate the social behavioral profile and hierarchy establishment in a semi-naturalistic environment to push our understanding of behavioral phenotyping related to stress exposure (Shemesh and Chen, 2023). Forkosh et al., (Forkosh et al., 2019) used an automated behavioral

monitoring system, called “Social Box” (SB), to investigate hierarchy rankings in mice utilizing a more naturalistic environment. This study uses homemade software to identify pose estimation and subsequently utilizes supervised behavioral analysis tools to calculate a set of social behaviors. The most important classifier for this study is the calculation of aggressive chases that is used, together with some other behaviors, to obtain David’s scores that reflect the hierarchy rankings between mice. The SB system can house up to four mice and can easily be utilized for longitudinal day and night recordings. Interestingly, in a follow-up study, the authors explore how social hierarchy rank can be used as a predictor of chronic stress exposure-outcome in a sex-dependent way, emphasizing the importance of investigating both sexes in the stress-related research field (Karamihalev et al., 2020). The data from the SB system exemplifies the enriched data output using longitudinal observations in a semi-naturalistic environment. This is extremely relevant as it allows for a more natural inference of motivation and decision-making processes crucial for hierarchy rankings. It would be interesting to see how the hierarchy rankings in the SB system compare to the classical approaches, especially in cohorts of animals that show altered motivation or decision-making. Another next step would be to integrate additional behavioral domains via implementing behavioral tasks during the longitudinal overnight recordings, which would be able to abolish the influences of initial arousal, sleep disturbances, and anxiety on the performance of the behavior. An example is the implementation of home cage operant conditioning devices, such as FED3, which allow the investigation of food intake, motivation, and reward processing (Nguyen et al., 2016). In order to boost the implementation of such a system for many different labs, the SB system should implement open-source pose-estimation and classifying software so that the identification of hierarchy in home-cage-like systems would easily be transferable.

A study that implemented a fully open-source software pipeline for stress phenotyping is Bordes et al., (Bordes et al., 2022). The study uses an open field environment with home-cage-like settings (using bedding material inside the maze) to explore chronic social defeat stress (CSDS)-induced social behavioral profiles. CSDS is a well-established stress model for investigating symptoms of depression, which ultimately leads to the avoidance of novel social stimuli in a subset of susceptible mice (Golden et al., 2011; Krishnan et al., 2007). However, the identification of social avoidance behavior is typically measured using the aforementioned social avoidance task or three-chamber task and thus relies on an oversimplified model of the social behavioral repertoire (Pearson et al., 2010; Bordes et al., 2022). In this study, an open-source python package DeepOpenField (DeepOF) was developed to analyze the stress effect using a set of supervised behavioral classifiers related to individualistic- and social behaviors, but also using an unsupervised clustering analysis of the individualistic and social behavioral interaction data. Both supervised and unsupervised pipelines showed a much stronger CSDS-induced difference in the free-social interaction task compared to the classical social avoidance task, indicating a more robust and clearer social behavioral profile using DeepOF. This study identified a subset of different social behavioral classifiers to be altered in stressed animals, which was subsequently used as an indicator for stress affectiveness by calculating a Z-score. Ultimately, the Z-score for stress affectiveness was significantly correlated with markers for stress physiology (end-point body weight and relative adrenal weight), which was not observed using the classical social avoidance IR, indicating that the Z-score analysis of DeepOF-identified behaviors provides a more complete social behavioral profile related to stress exposure than just the social avoidance IR. In addition, the mean-normalization of the data to the control group enables correcting for a potential batch effect. The DeepOF profiling ultimately provides a sensitive tool to assess individual stress resiliency and susceptibility. Interestingly, Ayash et al., (Ayash et al., 2023) observed that the social preference behavior of defeated C57/Bl6 mice in the three-chamber task is different between using a social conspecific from the aggressor strain (CD1) or a novel strain (e.g.,

129/Sv). In addition, they identified that defeated mice who do not avoid CD1 mice (classically labeled as resilient), show impaired learning in a fear conditioning model, and therefore the classification of resiliency is potentially incorrect. This is an interesting perspective and should be incorporated into the free social interaction task, to explore the social preference behavior after CSDS across different species.

Taken together, the increased information of data in video recognition using advanced pose estimation tools in a semi-naturalistic environment has expanded the understanding of the complexity of behavior. Implementations of these tools will lead to an increased understanding of the stress response system, for example, by enabling more ecologically relevant stressors, which are critical for the correct classification of resilient and susceptible animals (Lyons et al., 2023). In addition, these tools can dissect important influences of stress on behavior, such as the type of stressor (physiological versus psychological), stress intensity (chronic versus acute), and the dynamics of behavioral changes following a stressful challenge (Miranda et al., 2023).

4.3. Deep phenotyping benefits the integration of different factors influencing social behavior

Applying advanced tracking tools in semi-naturalistic environments offers the possibility to integrate different factors that shape social behavior, including stress exposure (stressed versus nonstressed), sex (female versus male), motivation (rewarding versus not rewarding), recording length (short versus longitudinal recordings), different developmental stages and aging (early life, adolescence, adulthood, older aged), and housing conditions (single housed, group-housed, or overcrowded), see Fig. 1. For example, the difference between short recordings (minutes to hours) and longitudinal recordings (day and night) is an important methodological consideration. Using shorter tests allows for high-throughput and more easily quantifiable data, and in addition is easily combined with other data modalities, such as neuronal activity data. However, these shorter tests are a snapshot during a certain time and are difficult to fully reflect non-arousal (baseline) behaviors. In contrast, longitudinal recordings are more suitable for investigating baseline behaviors and can be used for a truly free-choice behavioral task that is relying on no interaction with the experimenter. However, adding data modalities, such as neuronal activity is more difficult. Another important aspect is to distinguish between distinct developmental stages, which are known to differentially influence social behavior. First of all, there are specific vulnerable stages during life, in which animals are differentially influenced by environmental circumstances. For example, early life stress exposure influences the social behavioral profile and hierarchy establishment of adult mice in a sex-dependent manner (Bondar et al., 2018; Benner et al., 2014), but also the adolescent stage has been implicated as a critical stage for social behavioral development (Musardo et al., 2022; Endo et al., 2018). Secondly, aging is a strong factor influencing the social behavioral repertoire. Older aged animals have been found to show decreased explorative behaviors and increased grooming (Tran et al., 2021), while also a decreased number of social contacts and social novelty-seeking has been observed (Shoji et al., 2016; Shoji and Miyakawa, 2019). Moreover, the aggressive behavior towards conspecifics changes during aging as well, which is why for chronic social defeat stress only up to 6 months old aggressor mice can be used (Golden et al., 2011). A fundamental question for aging research is how behavior changes over longer periods of time (weeks to months). However, repeated behavioral testing with several classical behavioral tests influences the arousal and anxiety state of the animals, which ultimately leads to different behavioral outcomes between the tests (Voikar et al., 2004; Lad et al., 2010; McIlwain et al., 2001). As a consequence, researchers often use different cohorts of animals and test each cohort at a given age, making within-subject comparisons impossible, thereby increasing the risk of batch effects. A solution to this problem is to observe groups of animals in a semi-natural environment, for which no interference with anxiety,

or differences in the state of arousal can take place, as the animals are not in a specific test environment. Therefore, using advanced tracking tools in a semi-natural environmental set-up will not only facilitate fundamental research of social behavior, but also other research fields, such as stress and aging. The ability to test the same cohort of animals throughout the different stages of life will enable researchers to uncover unbiased behavioral trajectories over time. However, even though motion tracking data provides great insights into behavior, it does not provide the entire picture of the behavior profile. Currently, exciting developments are implementing different data modalities that can be measured in-vivo to provide an even richer understanding of the underlying behavior. The next section will discuss a set of data modalities that are currently being implemented in social behavioral analysis.

4.4. Integrating different behavioral and neuronal data modalities into social behavioral profiling

In recent years there has been an increased focus on the relevance of motion-tracking data for understanding behavior. This has been due to large breakthroughs in pose-estimation tools and their analysis, but also because video data is relatively easy and cheap to acquire. Furthermore, motion analysis can be easily interpreted, since researchers can map back their annotations and results to visual cues, which is significantly more difficult with other relevant types of data. However, behavior cannot be captured completely by motion tracking, as behavior is much more than just the posture of the animals. Therefore, new dimensions need to be included so that an even richer profile of the underlying behavior can be assessed. One option is to increase the data modalities using different sensory information of the animals, such as auditory data (e.g. mouse vocalizations), olfactory data (e.g. animal smell), and visual data (e.g. facial expressions) (Jabarin et al., 2022). Another option is to add physiological data modalities, for example, pupil dilation (Privitera et al., 2020), blood pressure, or heart rate, which can be used as a proxy for perceived stress (Buuse et al., 2001; Lemaire and Mormede, 1995; Farah et al., 2004; Swiercz et al., 2018). Lastly, adding neuronal data modalities will be important for future behavioral phenotyping, for example, by using neuronal activity (Padilla-Coreano et al., 2022), or neurotransmitter release data (Wu et al., 2022), which need to be wireless systems in order to sustain longitudinal recordings without restricting the animals (Fig. 3). The incorporation of different data modalities will increase the amount of data to be analyzed, which provides a new problem on how to tackle data integration. Open-source software is appearing to tackle these specific problems, such as CEBRA (Schneider et al., 2023). CEBRA aims to integrate data modalities to gain a more holistic unsupervised representation of behavior by enabling combined embeddings of motion tracking and neural activity data, which promises to facilitate the identification of the neural processes underlying motion. An interesting study by Padilla-Coreano et al., (Padilla-Coreano et al., 2022) utilized wireless electrophysiology devices to investigate and predict behavioral outcomes based on neuronal activity data. A different activity pattern was observed in the medial prefrontal cortex between social ranks and wins and losses in a reward competition task, which was visible 30 seconds prior to the start of the task. This study shows that the integration of neuronal activity data with behavior adds interesting insights to the understanding of behavior and the underlying mechanisms.

5. Conclusion

The history of the development of behavioral assessment has seen several influential methods that have their inherent advantages and disadvantages. Ethology applies observational research, which can study behavior in a descriptive way, but, among other problems, lacks the standardization of behavioral assessment and control over the environment. Comparative psychology methods reduced the complexity of the behavior by using a highly controllable environment but

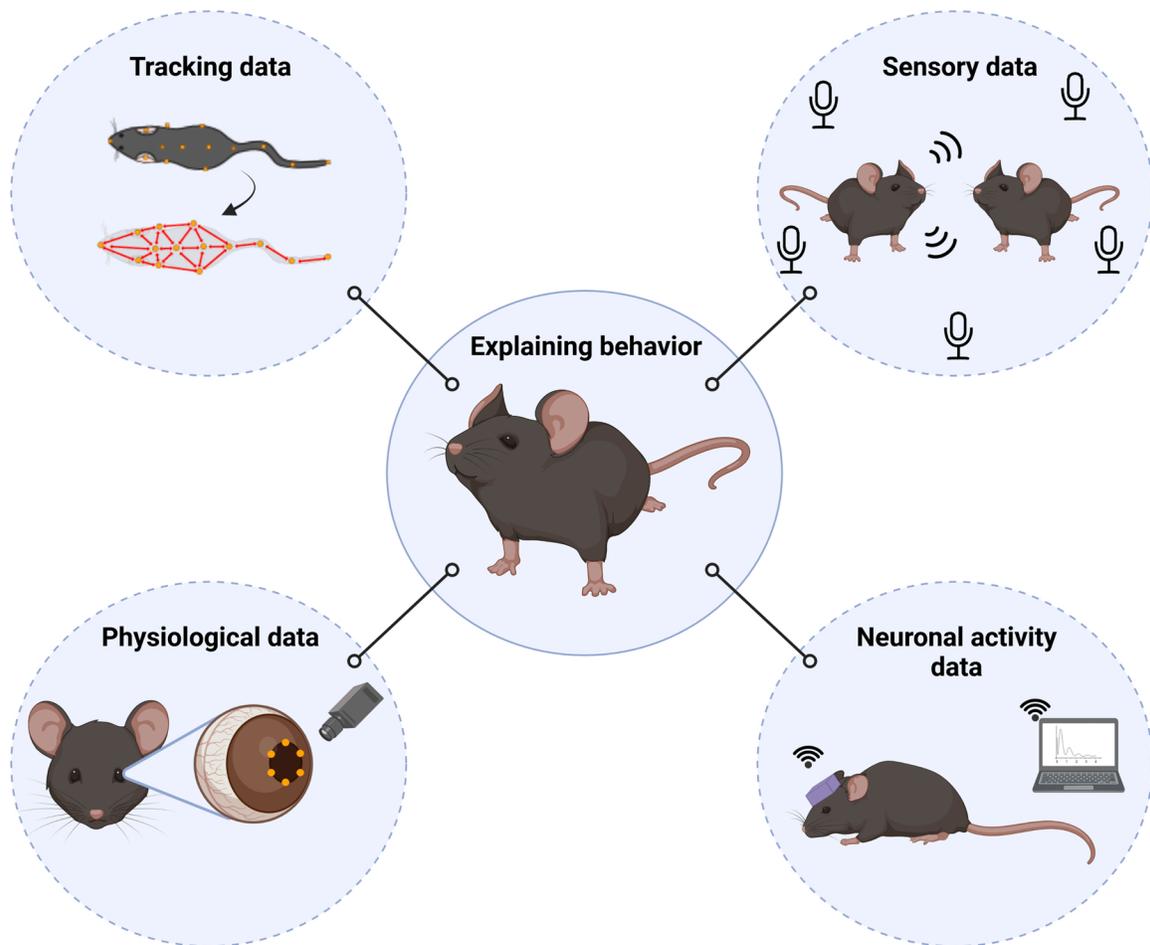


Fig. 3. Integrating data modalities in social behavioral profiling. Different data modalities need to be integrated into in-vivo recordings, in order to increase the understanding of rodent behavior. Diverse data modalities are currently being implemented, which include advanced tracking modalities using pose-estimation, sensory data (e.g. auditory, olfactory, and visual data), physiological data (e.g. pupil dilation, blood pressure or heart rate), and neuronal data modalities (e.g. neuronal activity data or neurotransmitter release).

oversimplified the underlying behavioral construct. The development of advanced and precise tracking tools, together with the numerous post-tracking analysis packages, has facilitated the (social) behavioral research field to assess behavior in a more naturalistic environmental set-up. This allows the investigation of behavior in a more descriptive way, without the interference of test-specific anxiety and arousal of the animals, but maintains a certain level of controllability over the environment. Further, the new behavioral analysis tools can be applied to classic behavioral tests to uncover and quantify novel behavioral traits. The implementation of body posture tracking and machine learning analysis will be beneficial to fundamental social behavioral research, but will also enable an increased understanding of the influences of many different factors that can influence social behavior, such as stress exposure, sex, motivation, recording length, different developmental stages, aging, and housing conditions. However, there are several challenges using such advanced tracking tools and an important step will be to develop an open-source centralized software system to allow a compatible ecosystem for tracking, annotation, and the embedding of multimodal behavioral data. Future research will increase the amount of data modalities, such as sensory, physiological, and neuronal activity data, and will thereby significantly enhance our understanding of the biological basis of social behavior and guide intervention strategies for behavioral abnormalities in psychiatric disorders.

Data Availability

No data was used for the research described in the article.

Acknowledgments

The authors thank London Aman for the language editing of the manuscript. Figures were created with BioRender.com. This work was supported by the Kids2Health grant of the Federal Ministry of Education and Research (01GL1743C) and the SAME-NeuroID grant of the European Research Executive Agency (101079181), and the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 813533.

Author contributions

The concept of the manuscript was established by JB and MVS. The writing of the manuscript by JB. Feedback and support by MVS, LM, and BMM.

References

- Aggleton, J.P., 1985. One-trial object recognition by rats. *Q. J. Exp. Physiol.* 37 (4), 279–294. <https://doi.org/10.1080/14640748508401171>.
- American Psychiatric Association, 2013. *Diagnostic and statistical manual of mental disorders*. Am. Psychiatr. Assoc. vol. 21.

- Ayash, S., Lingner, T., Ramisch, A., Ryu, S., Kalisch, R., Schmitt, U., Müller, M.B., 2023. Fear circuit-based neurobehavioral signatures mirror resilience to chronic social stress in mouse. *Proc. Nat. Acad. Sci.* 120 (17), e2205576120.
- Bariselli, S., et al., 2018. Role of VTA dopamine neurons and neurotrophin 3 in sociability traits related to nonfamiliar conspecific interaction. *Nat. Commun.* 9, 3173.
- Barnes, C.A., 1979. Memory deficits associated with senescence: a neurophysiological and behavioral study in the rat. *J. Comp. Physiol. Psychol.* 93, 74–104.
- Beery, A.K., Kaufner, D., 2015. Stress, social behavior, and resilience: insights from rodents. *Neurobiol. Stress* 1, 116–127.
- Benner, S., Endo, T., Endo, N., Kakeyama, M., Tohyama, C., 2014. Early deprivation induces competitive subordination in C57BL/6 male mice. *Physiol. Behav.* 137, 42–52.
- Berton, O., et al., 2006. Essential Role of BDNF in the mesolimbic dopamine pathway in social defeat stress (1979). *Science* 311, 864–868.
- Bidinosti, M., et al., 2016. CLK2 inhibition ameliorates autistic features associated with SHANK3 deficiency (1979). *Science* 351, 1199–1203.
- Blumstein, D.T., 2010. Towards an integrative understanding of social behavior: new models and new opportunities. *Front. Behav. Neurosci.* <https://doi.org/10.3389/fnbeh.2010.00034>.
- Bolhuis, J.J. & Giraldeau, L.-Alain. The study of animal behavior. (2005).
- Bolhuis, J.J., Giraldeau, L., Hogan, J.A., 2021. The study of animal behavior. in *The Behavior of Animals*, 2nd Edition. Wiley., pp. 1–11. <https://doi.org/10.1002/9781119109556.ch1>.
- Bondar, N.P., Lepeshko, A.A., Reshetnikov, V. v., 2018. Effects of early-life stress on social and anxiety-like behaviors in adult mice: sex-specific effects. *Behav. Neurol.* 1–13.
- Bordes, J., et al., 2022. Automatically annotated motion tracking identifies a distinct social behavioral profile following chronic social defeat stress, 2022.06.23.497350 [bioRxiv. https://doi.org/10.1101/2022.06.23.497350](https://doi.org/10.1101/2022.06.23.497350).
- Buuse, M., Acker, S.A.B.E., Flutterm, M., Kloet, E.R., 2001. Blood pressure, heart rate, and behavioral responses to psychological 'novelty' stress in freely moving rats. *Psychophysiology* 38, 490–499.
- de Chaumont, F., et al., 2019. Real-time analysis of the behaviour of groups of mice via a depth-sensing camera and machine learning. *Nat. Biomed. Eng.* 3, 930–942.
- Chen, P., Hong, W., 2018. Neural circuit mechanisms of social behavior. *Neuron* 98, 16–30.
- Chesler, E.J., Wilson, S.G., Lariviere, W.R., Rodriguez-Zas, S.L., Mogil, J.S., 2002. Influences of laboratory environment on behavior. *Nat. Neurosci.* 5, 1101–1102.
- Cordero, M., Sandi, C., 2007. Stress amplifies memory for social hierarchy. *Front. Neurosci.* 1, 175–184.
- Crawley, J.N., 1996. Unusual behavioral phenotypes of inbred mouse strains. *Trends Neurosci.* 19, 181–182.
- Darwin, C., 1859. On the Origin of Species by Means of Natural Selection, or Preservation of Favoured Races in the Struggle for Life, 1859. John Murray., London.
- Dodell-Feder, D., Tully, L.M., Hooker, C.I., 2015. Social impairment in schizophrenia. *Curr. Opin. Psychiatry* 28, 236–242.
- Donahue, R.J., Muschamp, J.W., Russo, S.J., Nestler, E.J., Carlezon, W.A., 2014. Effects of striatal ΔFosB overexpression and ketamine on social defeat stress–induced anhedonia in mice. *Biol. Psychiatry* 76, 550–558.
- Durand-de Cuttoli, R., et al., 2022. Distinct forms of regret linked to resilience versus susceptibility to stress are regulated by region-specific CREB function in mice. *Sci. Adv.* 8.
- Dwartz, M.F., Curley, J.P., Tye, K.M., Padilla-Coreano, N., 2022. Neural systems that facilitate the representation of social rank. *Philos. Trans. R. Soc. B: Biol. Sci.* 377.
- Ellis, L., 1995. Dominance and reproductive success among nonhuman animals: a cross-species comparison. *Ethol. Socio* 16, 257–333.
- Endo, N., et al., 2018. Multiple animal positioning system shows that socially-reared mice influence the social proximity of isolation-reared cagemates. *Commun. Biol.* 1, 225.
- Engelmann, M., Hädicke, J., Noack, J., 2011. Testing declarative memory in laboratory rats and mice using the nonconditioned social discrimination procedure. *Nat. Protoc.* 6, 1152–1162.
- Fan, Z., et al., 2019. Using the tube test to measure social hierarchy in mice. *Nat. Protoc.* 14, 819–831.
- Farah, V.M.A., Joaquim, L.F., Bernatova, I., Morris, M., 2004. Acute and chronic stress influence blood pressure variability in mice. *Physiol. Behav.* 83, 135–142.
- Fernández-Theoduloz, G., et al., 2019. Social avoidance in depression: a study using a social decision-making task. *J. Abnorm. Psychol.* 128, 234–244.
- Forkosh, O., et al., 2019. Identity domains capture individual differences from across the behavioral repertoire. *Nat. Neurosci.* 22, 2023–2028.
- Georgiou, P., et al., 2022. Experimenters' sex modulates mouse behaviors and neural responses to ketamine via corticotropin releasing factor. *Nat. Neurosci.* 25, 1191–1200.
- Gheusi, G., Bluthé, R.-M., Goodall, G., Dantzer, R., 1994. Social and individual recognition in rodents: methodological aspects and neurobiological bases. *Behav. Process.* 33, 59–87.
- Golden, S.A., Covington, H.E., Berton, O., Russo, S.J., 2011. A standardized protocol for repeated social defeat stress in mice. *Nat. Protoc.* 6, 1183–1191.
- Goodwin, N.L., Nilsson, S.R.O., Golden, S.A., 2020. Rage against the machine: advancing the study of aggression ethology via machine learning. *Psychopharmacol. (Berl.)* 237, 2569–2588.
- Goodwin, N.L., Nilsson, S.R.O., Choong, J.J., Golden, S.A., 2022. Toward the explainability, transparency, and universality of machine learning for behavioral classification in neuroscience. *Curr. Opin. Neurobiol.* 73, 102544.
- Graving, J.M., et al., 2019. DeepPoseKit, a software toolkit for fast and robust animal pose estimation using deep learning. *Elife* 8.
- Hänell, A., Marklund, N., 2014. Structured evaluation of rodent behavioral tests used in drug discovery research. *Front. Behav. Neurosci.* 8.
- Heimlich, J.E., Ardoin, N.M., 2008. Understanding behavior to understand behavior change: a literature review. *Environ. Educ. Res* 14, 215–237.
- Heuer, K., Rinck, M., Becker, E.S., 2007. Avoidance of emotional facial expressions in social anxiety: the approach–avoidance task. *Behav. Res. Ther.* 45, 2990–3001.
- Hsu, A.I., Yttri, E.A., 2021. B-SOiD, an open-source unsupervised algorithm for identification and fast prediction of behaviors. *Nat. Commun.* 12, 5188.
- Jabarin, R., Netser, S., Wagner, S., 2022. Beyond the three-chamber test: toward a multimodal and objective assessment of social behavior in rodents. *Mol. Autism* 13, 41.
- Jeon, D., Shin, H., 2011. A mouse model for observational fear learning and the empathetic response. *Curr. Protoc. Neurosci.* 57.
- Karamihalev, S., et al., 2020. Social dominance mediates behavioral adaptation to chronic stress in a sex-specific manner. *Elife* 9.
- Katz, S.J., Conway, C.C., Hammen, C.L., Brennan, P.A., Najman, J.M., 2011. Childhood social withdrawal, interpersonal impairment, and young adult depression: a mediational model. *J. Abnorm. Child Psychol.* 39, 1227–1238.
- Keum, S., Shin, H.-S., 2019. Neural basis of observational fear learning: a potential model of affective empathy. *Neuron* 104, 78–86.
- Krishnan, V., et al., 2007. Molecular adaptations underlying susceptibility and resistance to social defeat in brain reward regions. *Cell* 131, 391–404.
- Lad, H.V., et al., 2010. Behavioural battery testing: Evaluation and behavioural outcomes in 8 inbred mouse strains. *Physiol. Behav.* 99, 301–316.
- Larrieu, T., et al., 2017. Hierarchical status predicts behavioral vulnerability and nucleus accumbens metabolic profile following chronic social defeat stress. *e4 Curr. Biol.* 27, 2202–2210. e4.
- Lauer, J., et al., 2022. Multi-animal pose estimation, identification and tracking with DeepLabCut. *Nat. Methods* 19, 496–504.
- LeClair, K.B., et al., 2021. Individual history of winning and hierarchy landscape influence stress susceptibility in mice. *Elife* 10.
- Lemaire, V., Mormede, P., 1995. Telemetered recording of blood pressure and heart rate in different strains of rats during chronic social stress. *Physiol. Behav.* 58, 1181–1188.
- Li, L., et al., 2022. Social trauma engages lateral septum circuitry to occlude social reward. *Nature.* <https://doi.org/10.1038/s41586-022-05484-5>.
- Lindzey, G., Manosevitz, M., Winston, H., 1966. Social dominance in the mouse. *Psychon. Sci.* 5, 451–452.
- Lopez, J.P., et al., 2022. Ketamine exerts its sustained antidepressant effects via cell-type-specific regulation of Kcnq2. *Neuron.* <https://doi.org/10.1016/j.neuron.2022.05.001>.
- Lorsch, Z.S., et al., 2021. Computational analysis of multidimensional behavioral alterations after chronic social defeat stress. *Biol. Psychiatry* 89, 920–928.
- Lukas, M., et al., 2011. The neuropeptide oxytocin facilitates pro-social behavior and prevents social avoidance in rats and mice. *Neuropsychopharmacology* 36, 2159–2168.
- Luxem, K., et al., 2022. Identifying behavioral structure from deep variational embeddings of animal motion. *Commun. Biol.* 5, 1267.
- Lyons, D.M., Ayash, S., Schatzberg, A.F., Müller, M.B., 2023. Ecological validity of social defeat stressors in mouse models of vulnerability and resilience. *Neurosci. Biobehav. Rev.* 145, 105032.
- Mackintosh, J.H., Grant, E.C., 1963. A comparison of the social postures of some common laboratory rodents. *Behaviour* 21, 246–259.
- Marks, M., et al., 2022. Deep-learning-based identification, tracking, pose estimation and behaviour classification of interacting primates and mice in complex environments. *Nat. Mach. Intell.* 4, 331–340.
- Mathis, A., et al., 2018. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* 21, 1281–1289.
- Mathis, M., et al., 2020. DLC2Kinematics. a Post-Deep. *Modul. kinematic Anal.* <https://doi.org/10.5281/ZENODO.6669074>.
- Mathis, M.W., Mathis, A., 2020. Deep learning tools for the measurement of animal behavior in neuroscience. *Curr. Opin. Neurobiol.* 60, 1–11.
- Matthews, G.A., et al., 2016. Dorsal raphe dopamine neurons represent the experience of social isolation. *Cell* 164, 617–631.
- McIlwain, K.L., Merriweather, M.Y., Yuva-Paylor, L.A., Paylor, R., 2001. The use of behavioral test batteries: effects of training history. *Physiol. Behav.* 73, 705–717.
- Merlot, E., Moze, E., Bartolomucci, A., Dantzer, R., Neveu, P.J., 2004. The rank assessed in a food competition test influences subsequent reactivity to immune and social challenges in mice. *Brain Behav. Immun.* 18, 468–475.
- Miranda, L., Bordes, J., Gasperoni, S., Lopez, J.P., 2023. Increasing resolution in stress neurobiology: from single cells to complex group behaviors. *Stress* 26.
- Morel, C., et al., 2022. Midbrain projection to the basolateral amygdala encodes anxiety-like but not depression-like behaviors. *Nat. Commun.* 13, 1532.
- Morris, R.G.M., 1981. Spatial localization does not require the presence of local cues. *Learn Motiv.* 12, 239–260.
- Murra, D., et al., 2022. Characterizing the behavioral and neuroendocrine features of susceptibility and resilience to social stress. *Neurobiol. Stress* 17, 100437.
- Musardo, S., Contestabile, A., Knoop, M., Baud, O., Bellone, C., 2022. Oxytocin neurons mediate the effect of social isolation via the VTA circuits. *Elife* 11.
- Nguyen, K.P., O'Neal, T.J., Bolonduro, O.A., White, E., Kravitz, A. v., 2016. Feeding experimentation device (FED): A flexible open-source device for measuring feeding behavior. *J. Neurosci. Methods* 267, 108–114.
- Nietlisbach, G., Maercker, A., 2009. Social cognition and interpersonal impairments in trauma survivors with PTSD. *J. Aggress. Maltreat Trauma* 18, 382–402.

- Nilsson, S.R.O., et al., 2020. Simple behavioral analysis (SimBA) – an open source toolkit for computer classification of complex social behaviors in experimental animals, 2020.04.19.049452 bioRxiv. <https://doi.org/10.1101/2020.04.19.049452>.
- Ottensbreit, N.D., Dobson, K.S., Quigley, L., 2014. An examination of avoidance in major depression in comparison to social anxiety disorder. *Behav. Res. Ther.* 56, 82–90.
- Padilla-Coreano, N., et al., 2022. Cortical ensembles orchestrate social competition through hypothalamic outputs. *Nature* 603, 667–671.
- Park, M.-J., Seo, B.A., Lee, B., Shin, H.-S., Kang, M.-G., 2018. Stress-induced changes in social dominance are scaled by AMPA-type glutamate receptor phosphorylation in the medial prefrontal cortex. *Sci. Rep.* 8, 15008.
- Pearson, B.L., Defensor, E.B., Blanchard, D.C., Blanchard, R.J., 2010. C57BL/6J mice fail to exhibit preference for social novelty in the three-chamber apparatus. *Behav. Brain Res.* 213, 189–194.
- Peleh, T., Ike, K.G.O., Wams, E.J., Lebois, E.P., Hengerer, B., 2019a. The reverse translation of a quantitative neuropsychiatric framework into preclinical studies: Focus on social interaction and behavior. *Neurosci. Biobehav. Rev.* 97, 96–111.
- Peleh, T., Bai, X., Kas, M.J.H., Hengerer, B., 2019b. RFID-supported video tracking for automated analysis of social behaviour in groups of mice. *J. Neurosci. Methods* 325, 108323.
- Pereira, T.D., et al., 2019. Fast animal pose estimation using deep neural networks. *Nat. Methods* 16, 117–125.
- Pereira, T.D., et al., 2022. SLEAP: a deep learning system for multi-animal pose tracking. *Nat. Methods* 19, 486–495.
- Pisansky, M.T., Hanson, L.R., Gottesman, I.I., Gewirtz, J.C., 2017. Oxytocin enhances observational fear in mice. *Nat. Commun.* 8, 2102.
- Pöhlmann, M.L., et al., 2018. Pharmacological modulation of the psychiatric risk factor FKBP51 alters efficiency of common antidepressant drugs. *Front. Behav. Neurosci.* 12.
- Privitera, M., et al., 2020. A complete pupillometry toolbox for real-time monitoring of locus coeruleus activity in rodents. *Nat. Protoc.* 15, 2301–2320.
- Qin, L., et al., 2018. Social deficits in Shank3-deficient mouse models of autism are rescued by histone deacetylase (HDAC) inhibition. *Nat. Neurosci.* 21, 564–575.
- Qu, C., Ligneul, R., van der Henst, J.-B., Dreher, J.-C., 2017. An integrative interdisciplinary perspective on social dominance hierarchies. *Trends Cogn. Sci.* 21, 893–908.
- Rapanelli, M., et al., 2021. Behavioral, circuitry, and molecular aberrations by region-specific deficiency of the high-risk autism gene *Cul3*. *Mol. Psychiatry* 26, 1491–1504.
- Reguilón, M.D., Ferrer-Pérez, C., Manzanedo, C., Miñarro, J., Rodríguez-Arias, M., 2021. Ethanol intake in male mice exposed to social defeat: Environmental enrichment potentiates resilience. *Neurobiol. Stress* 15, 100413.
- Rein, B., et al., 2021. Reversal of synaptic and behavioral deficits in a 16p11.2 duplication mouse model via restoration of the GABA synapse regulator *Npas4*. *Mol. Psychiatry* 26, 1967–1979.
- Rein, B., Ma, K., Yan, Z., 2020. A standardized social preference protocol for measuring social deficits in mouse models of autism. *Nat. Protoc.* 15, 3464–3477.
- Richter, S.H., 2017. Systematic heterogenization for better reproducibility in animal experimentation. *Lab Anim. (NY)* 46, 343–349.
- Richter, S.H., Garner, J.P., Würbel, H., 2009. Environmental standardization: cure or cause of poor reproducibility in animal experiments? *Nat. Methods* 6, 257–261.
- Rinck, M., et al., 2010. Social anxiety predicts avoidance behaviour in virtual encounters. *Cogn. Emot.* 24, 1269–1276.
- Sakaguchi, T., Iwasaki, S., Okada, M., Okamoto, K., Ikegaya, Y., 2018. Ethanol facilitates socially evoked memory recall in mice by recruiting pain-sensitive anterior cingulate cortical neurons. *Nat. Commun.* 9, 3526.
- Sapolsky, R.M., 2005. The influence of social hierarchy on primate health. *Science* 308, 648–652.
- Scheggia, D., et al., 2022. Reciprocal cortico-amygdala connections regulate prosocial and selfish choices in mice. *Nat. Neurosci.* 25, 1505–1518.
- Schneider, S., Lee, J.H., Mathis, M.W., 2023. Learnable latent embeddings for joint behavioural and neural analysis. *Nature* 1–9.
- Segalín, C., et al., 2021. The mouse action recognition system (MARS) software pipeline for automated analysis of social behaviors in mice. *Elife* 10.
- Shemesh, Y., et al., 2013. High-order social interactions in groups of mice. *Elife* 2.
- Shemesh, Y., Chen, A., 2023. A paradigm shift in translational psychiatry through rodent neuroethology. *Mol. Psychiatry*. <https://doi.org/10.1038/s41380-022-01913-z>.
- Shoji, H., Miyakawa, T., 2019. Age-related behavioral changes from young to old age in male mice of a C57BL/6J strain maintained under a genetic stability program. *Neuropsychopharmacol. Rep.* 39, 100–118.
- Shoji, H., Takao, K., Hattori, S., Miyakawa, T., 2016. Age-related changes in behavior in C57BL/6J mice from young adulthood to middle age. *Mol. Brain* 9, 11.
- Skinner, B.F., 1948. 'Superstition' in the pigeon. *J. Exp. Psychol.* 38, 168–172.
- Solié, C., et al., 2022a. Superior Colliculus to VTA pathway controls orienting response and influences social interaction in mice. *Nat. Commun.* 13, 817.
- Solié, C., Girard, B., Righetti, B., Tapparel, M., Bellone, C., 2022b. VTA dopamine neuron activity encodes social interaction and promotes reinforcement learning through social prediction error. *Nat. Neurosci.* 25, 86–97.
- Sorge, R.E., et al., 2014. Olfactory exposure to males, including men, causes stress and related analgesia in rodents. *Nat. Methods* 11, 629–632.
- Sturman, O., et al., 2020. Deep learning-based behavioral analysis reaches human accuracy and is capable of outperforming commercial solutions. *Neuropsychopharmacology* 45, 1942–1952.
- Sulloway, F.J., 1982. Darwin and his finches: the evolution of a legend. *J. Hist. Biol.* 15, 1–53.
- Sun, J.J., et al., 2021. The Multi-Agent Behavior Dataset. *Mouse Dyadic Soc. Interact.* <https://doi.org/10.48550/arxiv.2104.02710>.
- Swiercz, A.P., Seligowski, A. v., Park, J., Marvar, P.J., 2018. Extinction of fear memory attenuates conditioned cardiovascular fear reactivity. *Front. Behav. Neurosci.* 12.
- Tamashiro, K.L.K., Nguyen, M.M.N., Sakai, R.R., 2005. Social stress: from rodents to primates. *Front. Neuroendocr.* 26, 27–40.
- Thierry, B., 2010. Darwin as a student of behavior. *C. R. Biol.* 333, 188–196.
- Thorndike, E.L., 1898. Animal intelligence: an experimental study of the associative processes in animals. *Psychol. Rev.: Monogr. Suppl.* 2, i–109.
- Tinbergen, N., 1939. Why do birds behave as they do?(II). *Bird-lore* 41, 23–30.
- Tolman, E.C., 1948. Cognitive maps in rats and men. *Psychol. Rev.* 55, 189–208.
- Torquet, N., et al., 2018. Social interactions impact on the dopaminergic system and drive individuality. *Nat. Commun.* 9, 3081.
- Toth, I., Neumann, I.D., 2013. Animal models of social avoidance and social fear. *Cell Tissue Res* 354, 107–118.
- Toth, I., Neumann, I.D., Slattery, D.A., 2013. Social fear conditioning as an animal model of social anxiety disorder. *Curr. Protoc. Neurosci.* 63.
- Tran, T., et al., 2021. Male–female differences in the effects of age on performance measures recorded for 23 hours in mice. *J. Gerontol.: Ser. A* 76, 2141–2146.
- Ujita, W., Kohyama-Koganeya, A., Endo, N., Saito, T., Oyama, H., 2018. Mice lacking a functional NMDA receptor exhibit social subordination in a group-housed environment. *FEBS J.* 285, 188–196.
- Voelkl, B., et al., 2020. Reproducibility of animal research in light of biological variation. *Nat. Rev. Neurosci.* 21, 384–393.
- Voikar, V., Vasar, E., Rauvala, H., 2004. Behavioral alterations induced by repeated testing in C57BL/6J and 129S2/Sv mice: implications for phenotyping screens. *Genes Brain Behav.* 3, 27–38.
- von Ziegler, L., Sturman, O., Bohacek, J., 2021. Big behavior: challenges and opportunities in a new era of deep behavior profiling. *Neuropsychopharmacology* 46, 33–44.
- Wahlsten, D., Metten, P., Crabbe, J.C., 2003. A rating scale for wildness and ease of handling laboratory mice: results for 21 inbred strains tested in two laboratories. *Genes Brain Behav.* 2, 71–79.
- Wang, F., et al., 2011. Bidirectional control of social hierarchy by synaptic efficacy in medial prefrontal cortex. *Science* 334, 693–697.
- Wang, F., Kessels, H.W., Hu, H., 2014. The mouse that roared: neural mechanisms of social hierarchy. *Trends Neurosci.* 37, 674–682.
- Weber, E.M., Dallaire, J.A., Gaskill, B.N., Pritchett-Corning, K.R., Garner, J.P., 2017. Aggression in group-housed laboratory mice: why can't we solve the problem? *Lab Anim.* 46, 157–161.
- Wei, D., Talwar, V., Lin, D., 2021. Neural circuits of social behaviors: innate yet flexible. *Neuron* 109, 1600–1620.
- Weinreb, C., et al., 2023. Keypoint-MoSeq: parsing behavior by linking point tracking to pose dynamics, 2023.03.16.532307 bioRxiv. <https://doi.org/10.1101/2023.03.16.532307>.
- Whishaw, I.Q., Dall, V.B., Kolb, B., 2006. Analysis of behavior in laboratory rats. in *The Laboratory Rat*. Elsevier, pp. 191–218. <https://doi.org/10.1016/B978-012074903-4/50011-X>.
- Wiltshcko, A.B., et al., 2020. Revealing the structure of pharmacobehavioral space through motion sequencing. *Nat. Neurosci.* 23, 1433–1443.
- Winslow, J.T., 2003. Mouse social recognition and preference. *Curr. Protoc. Neurosci.* 22.
- Wrenn, C.C., 2004. Social transmission of food preference in mice. *Curr. Protoc. Neurosci.* 28.
- Wu, Z., Lin, D., Li, Y., 2022. Pushing the frontiers: tools for monitoring neurotransmitters and neuromodulators. *Nat. Rev. Neurosci.* 23, 257–274.
- Ye, S., Mathis, A. & Mathis, M.W. Panoptic animal pose estimators are zero-shot performers. , 2022 doi:10.48550/arxiv.2203.07436.
- Yttri, E., Hsu, A., Schweihoff, J. & Schwarz, M. A-SOid, an active learning platform for expert-guided, data efficient discovery of behavior. (2023) doi:10.21203/rs.3.rs-2239160/v1.
- Zhou, T., et al., 1979. History of winning remodels thalamo-PFC circuit to reinforce social dominance. *Science* 357 (162–168), 2017.
- Zhou, T., Sandi, C., Hu, H., 2018. Advances in understanding neural mechanisms of social dominance. *Curr. Opin. Neurobiol.* 49, 99–107.
- Zimprich, A., et al., 2017. Assessing sociability, social memory, and pup retrieval in mice. *Curr. Protoc. Mouse Biol.* 7, 287–305.

Chapter 3

DeepOF: a python package for supervised and unsupervised pattern recognition in mice motion tracking data

Miranda, L., **Bordes, J.**, Pütz, B., Schmidt, M.V., Müller-Myhsok, B.

Originally published in:

The Journal of Open Source Software, 2023, 8(86), 5394

DeepOF: a Python package for supervised and unsupervised pattern recognition in mice motion tracking data

Lucas Miranda ¹, Joeri Bordes ², Benno Pütz ¹, Mathias V Schmidt ², and Bertram Müller-Myhsok ¹✉

¹ Research Group Statistical Genetics, Max Planck Institute of Psychiatry, Munich, Germany ² Research Group Neurobiology of Stress Resilience, Max Planck Institute of Psychiatry, Munich, Germany ✉ Corresponding author

DOI: [10.21105/joss.05394](https://doi.org/10.21105/joss.05394)

Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 

Editor: [Elizabeth DuPre](#)  

Reviewers:

- [@cellistigs](#)
- [@edeno](#)

Submitted: 12 April 2023

Published: 12 June 2023

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

Summary

DeepOF (Deep Open Field) is a Python package that provides a suite of tools for analyzing behavior in freely-moving rodents. Specifically, it focuses on postprocessing time-series data extracted from videos using [DeepLabCut](#) ([Mathis et al., 2018](#)). The software encompasses a diverse set of capabilities, such as:

- Loading DeepLabCut data into custom objects and incorporating metadata related to experimental design.
- Processing data, including smoothing, imputation, and feature extraction.
- Annotating behavioral motifs in a supervised manner, such as recognizing huddling and climbing, and detecting fundamental social interactions between animals.
- Embedding motion tracking data in an unsupervised manner using neural network models, which also facilitate end-to-end deep clustering.
- Conducting post-hoc analysis of results and visualization to compare patterns across animals under different experimental conditions.

The package is designed to work with various types of DeepLabCut input (single and multi-animal projects), includes comprehensive documentation, and offers interactive tutorials. Although many of its primary functionalities (particularly the supervised annotation pipeline) were developed with top-down mice videos in mind, tagged with a specific set of labels, most essential functions operate without constraints. As demonstrated in the accompanying scientific application paper ([Bordes et al., 2022](#)), DeepOF has the potential to enable systematic and thorough behavioral assessments in a wide range of preclinical research settings.

Statement of need

The field of behavioral research has experienced significant advancements in recent years, particularly in the quantification and analysis of animal behavior. Historically, behavioral quantification relied heavily on tests that were designed with either one or a few readouts in mind. However, the advent of deep learning for computer vision and the development of packages such as DeepLabCut, which enable pose estimation without the need for physical markers, have rapidly expanded the possibilities for non-invasive animal tracking ([Mathis et al., 2020](#)).

By transforming raw video footage into time series data of tracked body parts, these approaches have paved the way for the development of software packages capable of automatically

annotating behavior following a plethora of different approaches, increasing the number of patterns that can be studied per experiment with little burden on the experimenters.

For example, several tools offer options to detect predefined behaviors using supervised machine learning. Along these lines, programs like SimBA (Nilsson et al., 2020), MARS (Segalin et al., 2021), or TREBA (Sun et al., 2021), allow users to label a set of behaviors and train classifiers to detect them in new videos. They employ different labelling schemes which require different amounts of user input, and offer high flexibility in terms of the number of behaviors that can be detected. On the other hand, packages such as B-SOiD (Hsu & Yttri, 2021), VAME (Luxem et al., 2022), and Keypoint-MoSeq (Weinreb et al., 2023), aim for a more exploratory approach that does not require user labelling, but instead relies on unsupervised learning to segment time series into different behaviors. These packages are particularly useful when the user is interested in detecting novel behaviors, or when the number of behaviors is too large to be annotated manually. Moreover, some approaches have been developed to combine the best of both worlds, such as the the A-SOiD active learning framework (Schweihoff et al., 2022), and the semi-supervised DAART (Whiteway et al., 2021). While a thorough discussion on the advantages and disadvantages of each package is beyond the scope of this paper, further information can be found in this recent review (Bordes et al., 2023).

In contrast to other available options, DeepOF offers both supervised and unsupervised annotation pipelines, that allow researchers to test hypotheses regarding experimental conditions such as stress, gene mutations, and sex, in a flexible way (Figure 1).

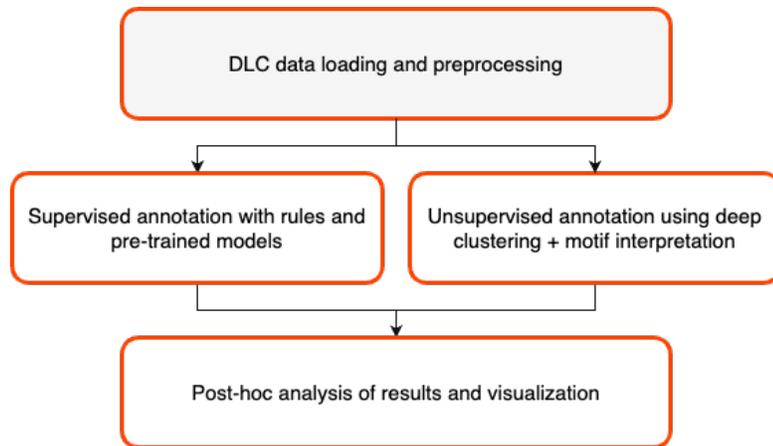


Figure 1: Scheme representing DeepOF workflow. Upon creating a project, DLC data can be loaded and preprocessed before annotating it with either a supervised pipeline (which uses a set of pre-trained models and rule-based annotators) or an unsupervised pipeline, which relies on custom deep clustering algorithms. Patterns retrieved with either pipeline can be passed to downstream post-hoc analysis tools and visualization functions.

The included supervised pipeline uses a series of rule-based annotators and pre-trained machine learning classifiers to detect when each animal is displaying a set of pre-defined behavioral motifs. The unsupervised workflow uses state-of-the-art deep clustering models to extract novel motifs without prior definition. DeepOF then provides an interpretability pipeline to explore what these retrieved clusters are in terms of behavior, which uses both Shapley Additive Explanations (SHAP) (Goodwin et al., 2022) and direct mappings from clusters to video. Moreover, regardless of whether the user chose the supervised annotation pipeline, the unsupervised one, or both, DeepOF provides an extensive set of post-hoc analysis and visualization tools.

When it comes to comparing it to other individual packages that use supervised and unsupervised

annotation, DeepOF stands out in several ways. First of all, it is the first package, to the best of our knowledge, to offer both options. Second, the supervised pipeline in DeepOF follows an opinionated philosophy, in the sense that it provides a set of pre-trained models that cannot be customized, but do not require user labels. This trades flexibility for ease of use, aiming at being a quick exploratory tool that can provide information on key individual and social behaviors with just a few commands. Furthermore, when it comes to the unsupervised pipeline, DeepOF provides three custom deep clustering algorithms capable of segmenting the behavioral time series, as well as the aforementioned built-in interpretability pipeline. If a user runs both pipelines, supervised annotations can be incorporated into this interpretability pipeline in quite a unique way, to detect associations between supervised and unsupervised patterns.

All in all, DeepOF is a comprehensive, end-to-end tool designed to transform DeepLabCut output into relatively quick, exploratory insights on behavioral shifts between experimental conditions, and pinpoint which behaviors are driving them.

Related literature

The DeepOF package has been used to characterize differences in behavior associated with Chronic Social Defeat Stress (CSDS) in mice, as presented in our preprint (currently in revision at the time of writing (Bordes et al., 2022)). There are several other ongoing projects involving the software, although none of them are published to this date.

Acknowledgements

We acknowledge contributions from Felix Agakov and Karsten Borgwardt.

Funding

This project has received funding from the European Union's Framework Programme for Research and Innovation Horizon 2020 (2014-2020) under the Marie Skłodowska-Curie Grant Agreement No. 813533-MSCA-ITN-2018.

References

- Bordes, J., Miranda, L., Müller-Myhsok, B., & Schmidt, M. V. (2023). Advancing social behavioral neuroscience by integrating ethology and comparative psychology methods through machine learning. *Neuroscience & Biobehavioral Reviews*, *151*, 105243. <https://doi.org/10.1016/J.NEUBIOREV.2023.105243>
- Bordes, J., Miranda, L., Reinhardt, M., Brix, L. M., Doeselaar, L. van, Engelhardt, C., Pütz, B., Agakov, F., Müller-Myhsok, B., & Schmidt, M. V. (2022). Automatically annotated motion tracking identifies a distinct social behavioral profile following chronic social defeat stress. *bioRxiv*. <https://doi.org/10.1101/2022.06.23.497350>
- Goodwin, N. L., Nilsson, S. R. O., Choong, J. J., & Golden, S. A. (2022). Toward the explainability, transparency, and universality of machine learning for behavioral classification in neuroscience. *Current Opinion in Neurobiology*, *73*, 102544. <https://doi.org/10.1016/j.conb.2022.102544>
- Hsu, A. I., & Yttri, E. A. (2021). B-SOiD, an open-source unsupervised algorithm for identification and fast prediction of behaviors. *Nature Communications*, *12*(1). <https://doi.org/10.1038/s41467-021-25420-x>
- Luxem, K., Mocellin, P., Fuhrmann, F., Kürsch, J., Miller, S. R., Palop, J. J., Remy, S., & Bauer, P. (2022). Identifying behavioral structure from deep variational embeddings of

- animal motion. *Communications Biology* 2022 5:1, 5(1), 1–15. <https://doi.org/10.1038/s42003-022-04080-7>
- Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience* 2018 21:9, 21(9), 1281–1289. <https://doi.org/10.1038/s41593-018-0209-y>
- Mathis, A., Schneider, S., Lauer, J., & Mathis, M. W. (2020). A Primer on Motion Capture with Deep Learning: Principles, Pitfalls, and Perspectives. *Neuron*, 108(1), 44–65. <https://doi.org/10.1016/j.neuron.2020.09.017>
- Nilsson, S. R., Goodwin, N. L., Choong, J. J., Hwang, S., Wright, H. R., Norville, Z. C., Tong, X., Lin, D., Bentzley, B. S., Eshel, N., McLaughlin, R. J., & Golden, S. A. (2020). Simple Behavioral Analysis (SimBA) – an open source toolkit for computer classification of complex social behaviors in experimental animals. *bioRxiv*. <https://doi.org/10.1101/2020.04.19.049452>
- Schweihoff, J. F., Hsu, A. I., Schwarz, M. K., & Yttri, E. A. (2022). A-SOiD, an active learning platform for expert-guided, data efficient discovery of behavior. *bioRxiv*. <https://doi.org/10.1101/2022.11.04.515138>
- Segalin, C., Williams, J., Karigo, T., Hui, M., Zelikowsky, M., Sun, J. J., Perona, P., Anderson, D. J., & Kennedy, A. (2021). The Mouse Action Recognition System (MARS) software pipeline for automated analysis of social behaviors in mice. *eLife*, 10, e63720. <https://doi.org/10.7554/eLife.63720>
- Sun, J. J., Kennedy, A., Zhan, E., Anderson, D. J., Yue, Y., & Perona, P. (2021, June). Task Programming: Learning Data Efficient Behavior Representations. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. <https://doi.org/10.1109/cvpr46437.2021.00290>
- Weinreb, C., Osman, M. A. M., Zhang, L., Lin, S., Pearl, J., Annapragada, S., Conlin, E., Gillis, W. F., Jay, M., Ye, S., Mathis, A., Mathis, M. W., Pereira, T., Linderman, S. W., & Datta, S. R. (2023). Keypoint-MoSeq: parsing behavior by linking point tracking to pose dynamics. *bioRxiv*. <https://doi.org/10.1101/2023.03.16.532307>
- Whiteway, M. R., Schaffer, E. S., Wu, A., Buchanan, E. K., Onder, O. F., Mishra, N., & Paninski, L. (2021). Semi-supervised sequence modeling for improved behavioral segmentation. *bioRxiv*. <https://doi.org/10.1101/2021.06.16.448685>

Chapter 4

Automatically annotated motion tracking identifies a distinct social behavioral profile following chronic social defeat stress

Bordes, J.*, Miranda, L.*, Reinhardt, M., Narayan, S., Hartmann, J., Newman, E.L., Brix, L.M., van Doeselaar, L., Engelhardt, C., Dillmann, L., Mitra, S., Ressler, K.J., Pütz, B., Agakov, F., Müller-Myhsok, B., Schmidt, M.V.

*Shared authorship

Originally published in:
Nature Communications

Automatically annotated motion tracking identifies a distinct social behavioral profile following chronic social defeat stress

Received: 19 September 2022

Accepted: 7 July 2023

Published online: 18 July 2023

 Check for updates

Joeri Bordes ^{1,6}, Lucas Miranda ^{2,3,6}, Maya Reinhardt ¹, Sowmya Narayan^{1,3}, Jakob Hartmann ⁴, Emily L. Newman ⁴, Lea Maria Brix^{1,3}, Lotte van Doeselaar ^{1,3}, Clara Engelhardt¹, Larissa Dillmann¹, Shiladitya Mitra¹, Kerry J. Ressler ⁴, Benno Pütz², Felix Agakov⁵, Bertram Müller-Myhsok ² ✉ & Mathias V. Schmidt ¹ ✉

Severe stress exposure increases the risk of stress-related disorders such as major depressive disorder (MDD). An essential characteristic of MDD is the impairment of social functioning and lack of social motivation. Chronic social defeat stress is an established animal model for MDD research, which induces a cascade of physiological and behavioral changes. Current markerless pose estimation tools allow for more complex and naturalistic behavioral tests. Here, we introduce the open-source tool DeepOF to investigate the individual and social behavioral profile in mice by providing supervised and unsupervised pipelines using DeepLabCut-annotated pose estimation data. Applying this tool to chronic social defeat in male mice, the DeepOF supervised and unsupervised pipelines detect a distinct stress-induced social behavioral pattern, which was particularly observed at the beginning of a novel social encounter and fades with time due to habituation. In addition, while the classical social avoidance task does identify the stress-induced social behavioral differences, both DeepOF behavioral pipelines provide a clearer and more detailed profile. Moreover, DeepOF aims to facilitate reproducibility and unification of behavioral classification by providing an open-source tool, which can advance the study of rodent individual and social behavior, thereby enabling biological insights and, for example, subsequent drug development for psychiatric disorders.

Stress is an essential aspect of our daily lives, which contributes to our mood and motivation. However, exposure to severe stress can have negative consequences and has become an increasing burden on society. In particular, stress-related disorders, such as major depressive disorder (MDD), have been steadily on the rise for the last decade¹. Our understanding of the behavioral and neurobiological mechanisms

related to MDD is limited, which is part of the reason for the only moderate success of current drug treatments². MDD is a complex and heterogeneous disorder, and its classification is dependent on a widespread set of symptoms. An important characteristic of MDD is the impairment of social functioning and lack of social motivation, which can lead to social withdrawal from society in extreme cases³. In

¹Research Group Neurobiology of Stress Resilience, Max Planck Institute of Psychiatry, 80804 Munich, Germany. ²Research Group Statistical Genetics, Max Planck Institute of Psychiatry, 80804 Munich, Germany. ³International Max Planck Research School for Translational Psychiatry (IMPRS-TP), 80804 Munich, Germany. ⁴Department of Psychiatry, Harvard Medical School, McLean Hospital, Belmont, MA 02478, USA. ⁵Pharmatics Limited, Edinburgh EH16 4UX Scotland, UK. ⁶These authors contributed equally: Joeri Bordes, Lucas Miranda. ✉e-mail: bmm@psych.mpg.de; mschmidt@psych.mpg.de

addition, disturbances in social behavior are an important risk factor for developing MDD, as poor social networks are linked to lowered mental and physical health^{4,5}. The impact of social interactions was highlighted during the COVID-19 pandemic, where a substantial part of society experienced little to no social interactions for a sustained period. An increasing number of studies are now reporting the enormous impact of the pandemic, emphasizing a dramatic increase in the prevalence of stress-related disorders, in particular MDD^{6,7}. Unfortunately, there is still a lack of awareness of the importance of social interactions and their role in stress-related disorders. Therefore, it is crucial to increase the understanding of the biological and psychological mechanisms behind MDD, and the influence of social behavior on the development of MDD.

Along these lines, animal models have an important role in MDD research. Although unable to recreate the exact nature of the disorder in humans, they provide a controlled environment where symptoms of MDD can be investigated^{8,9}. The well-established chronic social defeat stress (CSDS) paradigm is continuously used for studying symptoms of MDD in animals^{10,11}. In the CSDS model, mice are subjected daily to severe physical and non-physical stressors from aggressive mice for several weeks, which results in the chronic activation of the physiological stress response system, leading to bodyweight differences, enlarged adrenals, and elevated levels of corticosterone¹². In addition, animals subjected to CSDS show stress-related behaviors such as social avoidance, anhedonia, reduced goal-directed motivation, and anxiety-like behavior^{10,13–16}. Especially CSDS-induced social avoidance behavior, which is the avoidance of a novel conspecific, is a recognized phenomenon that is used to investigate the social neurobiological mechanisms related to chronic stress exposure and stress-related disorders^{11,17,18}.

Currently, several social behavioral tasks can assess different constructs of social behavior, particularly the social avoidance task¹⁸. It is important that these behavioral tasks are conducted with control over the environment to investigate the effects of external stimuli, such as stress exposure. For decades there has been a trend to standardize and simplify these tests to allow for greater comparability and higher throughput. Unfortunately, this has led to an oversimplification of the social behavioral repertoire and increased the risk for cross-over effects by other types of behavior, such as anxiety-related behavior. Moreover, due to limitations in tracking software, the analysis of the interaction between multiple freely moving animals remained difficult, which further limited the complexity of the behavioral assessment. Social behavior is a complex behavioral construct, which relies on many different types of behavioral interactions, that often are too complicated, time-intensive, and repetitive to assess manually^{19–21}. Ultimately, this can lead to poor reproducibility of the social behavioral construct, as observed for social approach behavior²².

The current advancement in automatically annotated behavioral assessment, however, allows for high-throughput analysis using pose estimation, involving both supervised classification (intending to extract pre-defined and characterized traits) and unsupervised clustering (which aims to explore the data and extract patterns without external information)^{23–28}. Importantly, the open-source tool DeepLabCut has provided a robust and easily accessible system for deep-learning-based motion tracking and markerless pose estimation^{29,30}. The use of supervised classification, by defining the behavioral patterns of interest a priori, is a powerful tool that simplifies the analysis by using predefined relevant behavioral constructs without losing the complexity of social behavior. Furthermore, recent studies have shown the value of unsupervised clustering in addition to a supervised analysis, which can reveal novel and more complex structures of behavior^{19,26,31–33}. By acting in a more exploratory fashion, these practices can not only assist the discovery of novel traits but also direct researchers toward the main behavioral axes of variation across cohorts of interest. In addition, both the supervised and unsupervised

analysis approaches can provide more transparency for the behavioral definition and can easily be shared via online repositories, which contributes to a more streamlined definition of behavior across different labs^{21,34}. These computational tools can elevate the current understanding of the influences of stress exposure on behavior, by increasing the resolution of the observed behavioral output³⁵.

Therefore, the current study provides an application of our open-source tool DeepOF³⁶, which enables users to delve into the individual and social behavioral profiles of mice using DeepLabCut-annotated pose estimation data (Fig. 1). DeepOF provides two main workflows; a supervised behavioral analysis pipeline, which applies a set of annotators and pre-trained classifiers to detect defined individual and social traits, and an unsupervised analysis pipeline, capable of embedding the motion-tracking data of one or more animals in a latent behavioral space, pointing toward differences across experimental conditions without any label priming. Furthermore, DeepOF can retrieve unsupervised clusters of behavior that can be compared across conditions and therefore hint at previously unrecognized behavioral patterns that trigger new hypotheses. We describe a distinct social behavioral profile following CSDS in mice that can be recapitulated with both supervised and unsupervised workflows. Moreover, the current study observes a clear state of arousal upon exposure to a novel social conspecific that fades over time, which provides crucial insights for the quantification of optimal behavioral differences across time and experimental conditions.

Results

The supervised pipeline provided by DeepOF yields generalizable annotations

As expected, all rule-based behaviors show high performance when compared to manual labeling, which constitutes an argument in favor of simple behavioral tagging (Supplementary Fig. 1).

When evaluating the performance of the huddle classifier, balanced accuracy in the training set (0.78 ± 0.005) was marginally higher than in both validation settings (suggesting no overfitting), and performance on the internal validation (0.75 ± 0.046) was not significantly higher than performance on the external validation (0.75 ± 0.04) suggesting excellent generalization to new datasets (independent samples t-test: $T(7.34) = -0.03$, $p = 0.51$, Supplementary Fig. 2A). In addition, pseudo-labeling conducted on the external dataset showed a strong and significant correlation between total behavior duration across manual and predicted labels (Supplementary Fig. 2B). Finally, the SHAP analysis of the deployed classifier revealed low head movement, low spine stretch, low body area, and low locomotion speed as the most important features of the model, which goes in line with the accepted definition of the behavior (Supplementary Fig. 2C).

The physiological and behavioral hallmarks of stress are reproduced by CSDS

The CSDS paradigm was performed to maintain stress exposure for several weeks (Fig. 2A), which induced dysregulation of the hypothalamic-pituitary-adrenal axis (HPA-axis) and a stress-related behavioral profile. Male mice that were subjected to CSDS showed clear hallmarks of stress exposure, as observed by a significant increase in body weight during the stress paradigm, which was especially apparent towards the end of the stress (Fig. 2B, C), an increase in relative adrenal weight (Fig. 2D), reduced locomotion and time spent in the inner zone of the OF (Fig. 2E, F), and a significantly reduced SA-ratio in the SA task (Fig. 2G). Notably, no bodyweight difference was observed at the beginning of the CSDS paradigm (Fig. 2B).

Further exploration of the OF data using PCA across four 2.5 min consecutive time bins showed that all time bins were significantly different from each other, suggesting that they all should be included in further behavioral analysis of the OF data (Supplementary Fig. 3A,

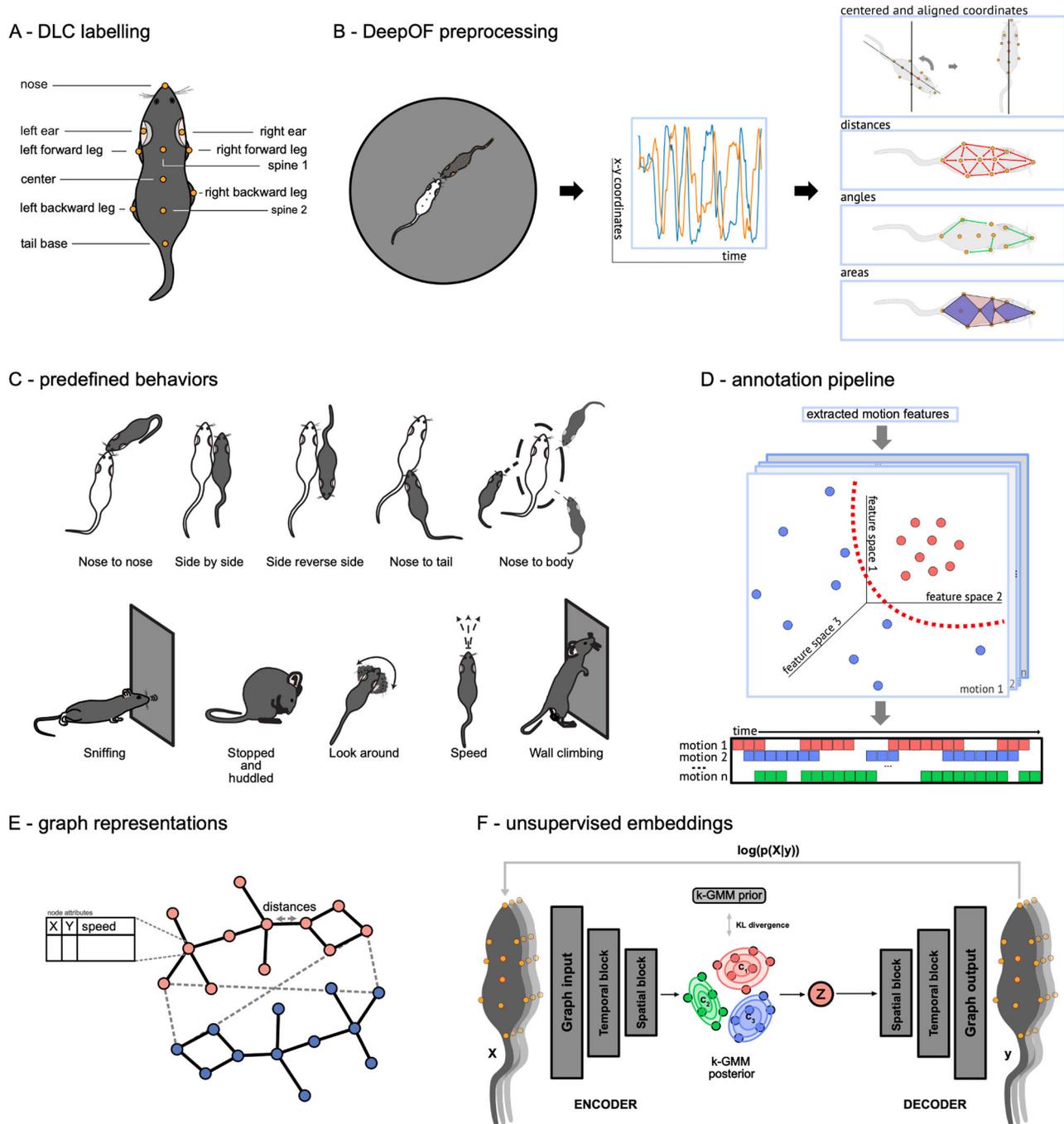


Fig. 1 | DeepOF workflow. **A** 11 labels were tagged on each annotated mouse using DeepLabCut. **B** DeepOF preprocessing pipeline. One or two mice (a C57Bl/6N experimental subject and a CDI social companion depending on the dataset) were tagged using the provided DeepLabCut models. After tracking body parts with DeepLabCut, DeepOF was used to smooth the retrieved trajectories, interpolate outliers, and extract features (including coordinates, distances, angles, areas, speeds and accelerations). **C** Set of predefined behaviors that the DeepOF supervised pipeline can retrieve. These include dyadic motifs (such as nose-to-nose contacts) and individual motifs (such as climbing), which are reported individually for all tracked mice. The stopped-and-huddled classifier²⁸ is abbreviated as “huddle” in DeepOF output (not to be confused with group huddling behavior⁶⁷). **D** Schematic representation of the supervised pipeline in DeepOF. A set of extracted motion features (only three dimensions are shown for visualization purposes) are fed to a set of rule-based annotators and pre-trained classifiers, which report the presence of each behavioral trait at each time by learning how the corresponding trait is distributed in the feature space (red dots). The set of

classifiers then yields a table indicating the presence of each motif across time, which can be used for further analysis. Note that annotators are not necessarily mutually exclusive, as several predictors can be triggered at the same time. **E** Graph representation of animal trajectories used by DeepOF in the unsupervised pipeline. All 11 body parts per animal are connected using a pre-designed (but customizable) adjacency matrix. Nodes are annotated with x , y coordinates and speed of each body part at each given time, and edges with the corresponding distances. This representation can also handle multi-animal settings, where the graphs of individual animals are connected with nose-to-nose, nose-to-tail, and tail-to-tail edges. **F** Schematic representation of the deep neural network architecture used for the unsupervised clustering of behavior. Data is embedded with a sequence-aware spatio-temporal graph encoder, and clustered at the same time by selecting the argmax of the likelihood of the components of a mixture-of-Gaussians latent posterior. Unidirectional black arrows indicate forward propagation, and gray arrows indicate the reconstruction and KL divergence terms of the loss function, the latter of which minimizes the distance to an also mixture-of-Gaussians prior.

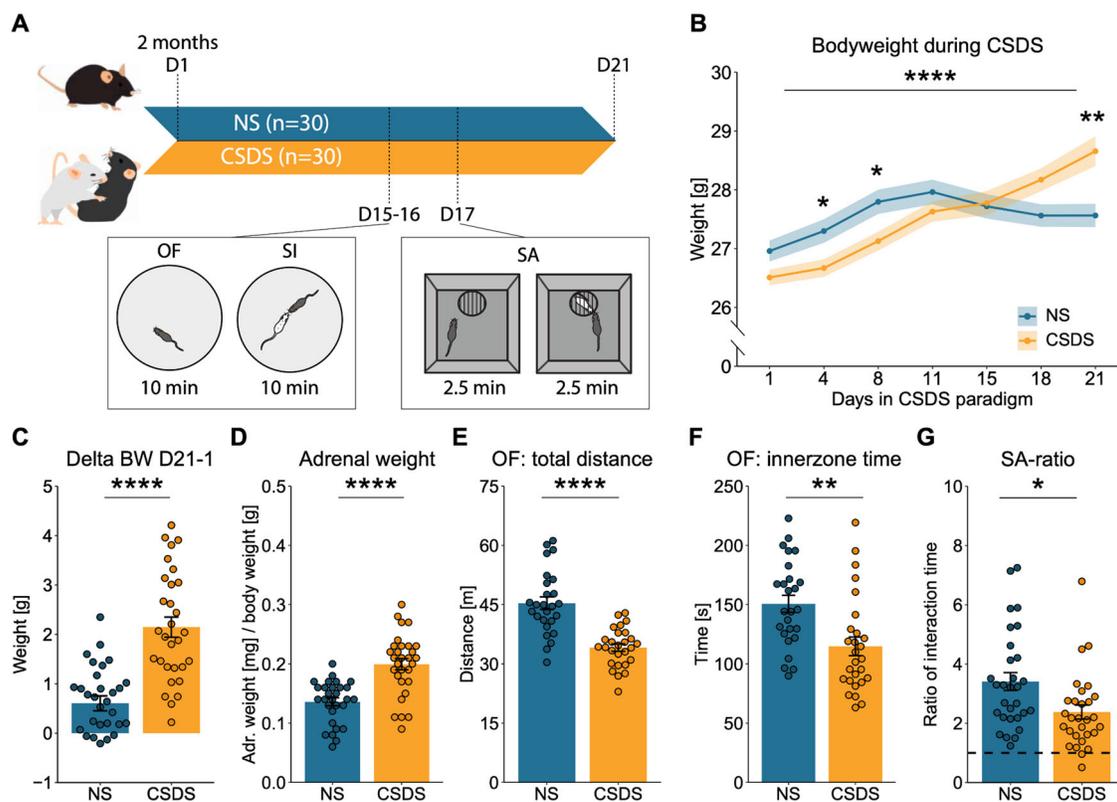


Fig. 2 | Classical hallmarks for chronic social defeat stress. **A** Experimental timeline for the CSDS paradigm and behavioral testing, including the open field (OF) and social interaction (SI) task on day 15–16 (animals were divided between the two days) and social avoidance (SA) task on day 17. **B** Significant increase of body weight after CSDS exposure (two-way ANOVA: within-subject effect of time: $F(6,406) = 13.58, p = 4.59 \times 10^{-14}$, as well as time \times condition interaction effect: $F(6,406) = 6.13, p = 3.65 \times 10^{-6}$, but no between-subject effect on condition: $F(1,406) = 0.20, p = 0.653$). Post-hoc analysis with Benjamini Hochberg revealed no significant difference on day 1, 11, 15, and 18, but there was a significant difference on day 4 ($T(1,58) = 6.36, p = 0.033$), 8 ($T(1,58) = 6.55, p = 0.033$), and 21 ($T(1,58) = 11.57, p = 0.007$). **C** The delta body weight during the CSDS paradigm (day 21–day 1) was

significantly increased in CSDS-exposed animals (Two-tailed independent samples t-test: $T(58) = -6.09, p = 9.8 \times 10^{-8}$). **D** Increase of relative adrenal weight after CSDS exposure (Two-tailed independent samples t-test: $T(57) = -5.44, p = 1.15 \times 10^{-6}$). **E** The total locomotion in the OF was reduced after CSDS exposure (Two-tailed independent samples t-test: $T(51) = 6.15, p = 1.18 \times 10^{-7}$). **F** The inner zone time in the OF was reduced after CSDS exposure (Two-tailed independent samples t-test: $T(51) = 3.37, p = 0.0015$). **G** The SA-ratio was reduced in the SA task after CSDS exposure (Two-tailed wilcoxon test: $W = 617, p = 0.006$). The timeline and bar graphs are presented as mean \pm standard error of the mean and all individual samples as points. $N = 30$ for NS and CSDS for (**B–G**). Source data are provided as a Source Data file.

B). The OF PCA between conditions revealed a significant difference and showed the importance of the OF parameters, in which total distance, look-around, and sniffing came out as the top contributing behaviors (Supplementary Fig. 3C, D). A significant stress effect was observed for the total distance, look-around, and inner-zone time throughout the different time bins, whereas sniffing was altered, but not in all time bins (Supplementary Fig. 3E–J). Importantly, even though a stress-induced effect can be found in the OF task, a general habituation effect to the OF in both NS and CSDS can be observed, as total distance reduces over time, while look-around and sniffing increase. The successful habituation to the novel environment is crucial for the subsequent SI task to allow full attention to the novel social conspecific (Supplementary Fig. 3E–G).

DeepOF social behavioral classifiers show a stronger PCA separation for stress exposure than social avoidance

The social behavioral pattern during the SI task was investigated in four non-overlapping time bins of 2.5 min each to match the time frame in the SA task. Principal component analysis (PCA) was performed to show the difference between time bins in the social behavioral profile regardless of the animal's stress condition (Fig. 3A). Interestingly, the PCA showed a significant effect between the time bins, in which the first 2.5 min time bin was significantly different from the subsequent ones (5, 7.5, and 10 min). In contrast, the subsequent time bins did not show variation between one another (Fig. 3B). This

suggests that the different time bins in the SI task are an important variable, and that the first 2.5 min time bin should be specifically investigated. Next, the SA and SI tasks were compared on their ability to distinguish between NS and CSDS animals. PCAs were performed for the SA task (Fig. 3C) and the 2.5 min time bin SI data (Fig. 3D, E), both of which showed a significant difference between the conditions in the principal component (PC) 1 eigenvalues (Fig. 3C–E). However, the SI task showed a clearer separation of the conditions than the SA task, suggesting that the SI task is a more powerful tool for identifying stressed animals than the SA task. In addition, the PC1 top contributing behaviors for the 2.5 min time bin SI data were calculated using the corresponding rotated loading scores (Fig. 3F). The top five contributing behaviors were reported as essential behaviors for identifying the stressed phenotype, which consisted of B-huddle, B-look-around, B-nose-to-tail, B-speed, and B-nose-to-body from the C57Bl/6N animal, whereas the other behaviors within the top 10 were either contributing to the CD1 animal or had a low rotated loading score (Fig. 3F). Here, “B-” indicates behaviors related to or initiated by the C57Bl/6N animals, whereas “W-” refers to the CD1.

DeepOF social behavioral classifiers are strongly altered by CSDS

Next, the influence of the CSDS on the top five contributing behaviors in the SI task was investigated. In accordance with the PCA time bin analysis, a clear stress-induced effect was observed, with elevated

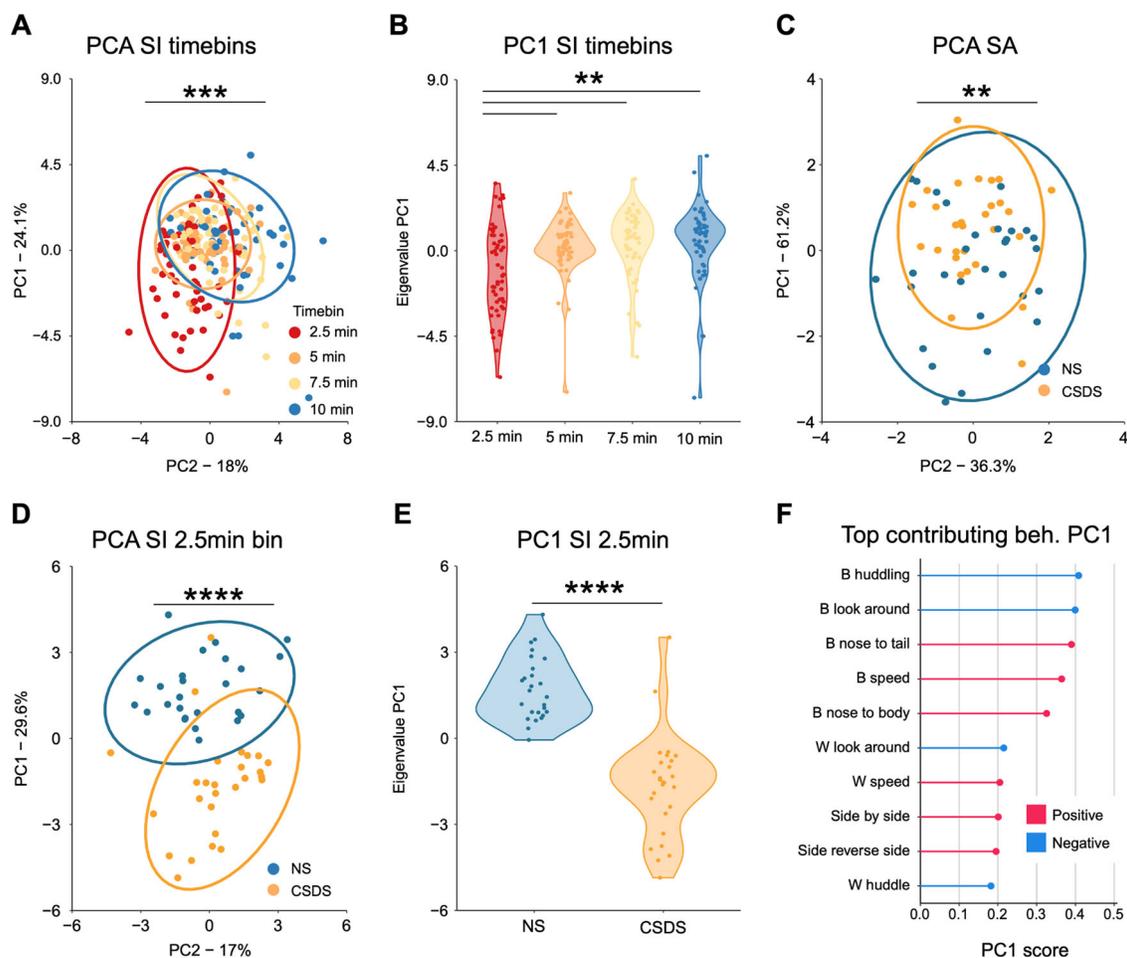


Fig. 3 | Social interaction binning yields more separable PCA projections than the social avoidance task. **A** In the SI data a PCA revealed that the first 2.5 min time bin is significantly different from the other time bins. (Kruskal-Wallis test: $H(3) = 19.90$, $p = 0.0002$. **B** The PC1 eigenvalues of the SI time bin PCA. Post-hoc Wilcoxon: 2.5 min vs. 5 min ($W = 957$, $p = 0.01$), 2.5 min vs. 7.5 min ($W = 860$, $p = 0.0018$), 2.5 min vs. 10 min ($W = 811$, $p = 0.0011$). **C** The SA task PCA showed a significant difference in the PC1 eigenvalues between conditions. The PCA data consisted of the SA-ratio, total time spent with the non-social stimulus, and total time spent with the social stimulus. Two-tailed independent samples t -test: $T(57) = -2.84$, $p = 0.006$. **D** The SI 2.5 min time bin PCA showed a significant difference in the PC1 eigenvalues between conditions. The PCA data consisted of all the SI DeepOF behavioral classifiers, as listed in Fig. 1C. Two-tailed independent

samples t -test: $T(51) = 8.28$, $p = 5.39e-11$. **E** The PC1 eigenvalues of the 2.5 min time bin SI task. **F** The top contributing behaviors of the SI 2.5 min time bin in PC1 using the corresponding rotated loading scores. The top five behaviors were reported as the essential behaviors for identifying stress exposure (B-huddle (-0.41), B-look-around (-0.40), B-nose-to-tail (0.39), B-speed (0.36), B-nose-to-body (0.33). “B-” indicates C57Bl/6N behaviors and “W-” indicates CD1 behaviors. The PCA graphs (Fig. 3A, C, D) are provided with a 95% confidence ellipse and all individual samples as points. Further PC1 analyses (Fig. B, E) are represented with a violin plot and all individual samples as points. In Fig. 3F the absolute score of the PC1 value is represented by the point. $N = 26$ for NS and $n = 27$ for CSDS in (A, B, D–F) and $n = 30$ for NS and CSDS in (C). Source data are provided as a Source Data file.

duration in the CSDS animals for B-look-around (Fig. 4A, B) and B-huddle (Fig. 4C, D), while lowered for the B-speed (Fig. 4E, F), B-nose-to-tail (Fig. 4G, H), and B-nose-to-body (Fig. 4I, J). The total duration per time bin for the top contributing behaviors showed the strongest CSDS-induced effect in the 2.5 min time bin data (supplemental Fig. 4, timeline graphs), compared to the 5, 7.5, and 10 min time bins. In addition, supplemental Fig. 4 shows the 10 min total duration and time bin analyses for all other DeepOF behavioral classifiers, in which a significant stress effect is observed for B-sniffing, B-wall-climbing, and Side-by-side.

Z-score for DeepOF social interaction correlates with Z-score for stress physiology

The Z-score of stress physiology was calculated using the relative adrenal weight and body weight on day 21 of the CSDS. The stress physiology Z-score provides a strong CSDS profiling tool and was used for correlation analysis between the SA and SI tasks. Even though the behavioral and physiological readouts were not obtained at the same

time, the former can be used as a proxy of the impact of the stress exposure, and are expected to be stable during the last week of the CSDS pipeline. No significant correlation was observed between the Z-score of stress physiology and the SA ratio (Fig. 5A). Subsequently, the Z-score of SI was calculated by using the 2.5 min time bin of the top five contributing behaviors in the SI task (Fig. 4). Stress physiology and SI Z-score showed a significant positive correlation (Fig. 5B), which indicates that the SI Z-score provides a stronger tool for CSDS profiling compared to the SA ratio. Next, correlation analyses were performed between the Z-score of SI and all other behavioral and physiological measurements which indicated a strong correlation with several OF parameters. Highly affected OF parameters, such as speed, distance, inner zone entries, and look-around might be directly related to social anxiety and warrant further investigation. Interestingly, no correlation with the SA ratio was observed (Fig. 5C).

Notably, the SA task is extensively used to distinguish resilient and susceptible animals in the CSDS paradigm^{10,17}, and depending on the protocol and stress severity this can give a distinction between

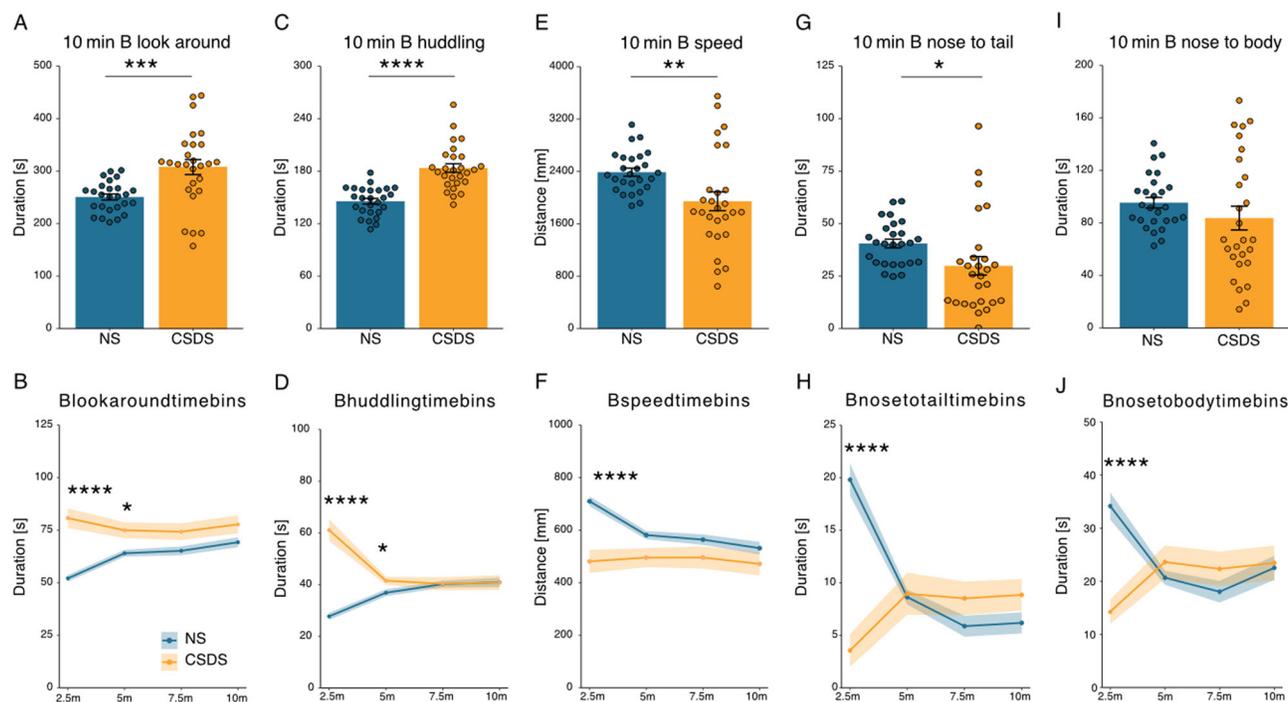


Fig. 4 | Top contributing behaviors in the social interaction task for 10 min total duration and time bins. **A** The total duration of B-look-around. Two-tailed Welch: $T(34.1) = -3.71$, $p = 0.0007$. **B** Time bin for B-look-around. Benjamini Hochberg (BH) posthoc for the 2.5 min time bin: ($T(51) = 33.46$, $p = 1.78e-6$) and the 5 min time bin ($T(51) = 6.84$, $p = 0.024$), but not for the 7.5 and 10 min time bins ($p = 0.067$, $p = 0.093$, respectively), two-way ANOVA: condition effect: $F(1,208) = 37.45$, $p = 4.59e-9$, time effect: $F(1,208) = 4.02$, $p = 0.046$, and condition \times time effect: $F(1,208) = 8.87$, $p = 0.003$. **C** The total duration of B-huddle. Two-tailed independent samples t -test: $T(51) = -6.40$, $p = 4.8e-8$. **D** Time bin for B-huddle. Wilcoxon posthoc for the 2.5 min time bin ($W(26,27) = 63.5$, $p = 1.3e-6$), and the 5 min time bin ($W(26,27) = 204$, $p = 0.018$), but not for the 7.5- and 10 min time bins ($p = 0.52$, $p = 0.52$, respectively), Kruskal-Wallis: 2.5 min: $p = 1.25e-6$, 5 min: $p = 0.018$, 7.5 min: $p = 0.51$, and 10 min: $p = 0.51$. **E** The total duration of B-speed. Two-tailed Welch: $T(35.04) = 2.84$, $p = 0.0074$. **F** Time bin for B-speed. BH posthoc for the 2.5 min time bin ($T(51) = 22.41$, $p = 7.16e-5$), but not for the 5-, 7.5-, and 10 min time bins

($p = 0.076$, $p = 0.20$, $p = 0.24$, respectively), two-way ANOVA: condition effect: $F(1,208) = 22.60$, $p = 3.72e-6$, time effect: $F(1,208) = 7.51$, $p = 0.007$, and condition \times time effect: $F(1,208) = 6.34$, $p = 0.013$. **G** The total duration of B-nose-to-tail. Two-tailed Welch: $T(36.70) = 2.18$, $p = 0.036$. **H** Time bin for B-nose-to-tail. Wilcoxon posthoc for the 2.5 min time bin ($W(26,27) = 660$, $p = 1.5e-7$), but not for the 5-, 7.5-, and 10 min time bins ($p = 0.19$, $p = 0.49$, $p = 0.49$, respectively), Kruskal-Wallis: 2.5 min: $p = 1.43e-7$, 5 min: $p = 0.18$, 7.5 min: $p = 0.48$, 10 min: $p = 0.48$. **I** The total duration of B-nose-to-body. Welch: $T(35.85) = 1.18$, $p = 0.24$. **J** Time bin for B-nose-to-body. Wilcoxon posthoc for the 2.5 min time bin ($W(26,27) = 626.5$, $p = 3.97e-6$), but not for the 5-, 7.5- and 10 min time bins ($p = 0.85$, $p = 0.85$, $p = 0.85$, respectively), Kruskal-Wallis: 2.5 min: $p = 3.8e-6$, 5 min: $p = 0.85$, 7.5 min: $p = 0.85$, 10 min: $p = 0.85$. The timeline and bar graphs are presented as mean \pm standard error of the mean and all individual samples as points. $N = 26$ for NS and $n = 27$ for CSDS in (A–J). Source data are provided as a Source Data file.

resilient and susceptible animals (Fig. 5D–F). Interestingly, while clearly differentiating affected and non-affected individuals, the DeepOF module does not find a distinction between SA-ratio-defined susceptibility and resiliency on the 2.5 min bin SI DeepOF behavioral classifiers (Fig. 5G–M), indicating that the DeepOF behavioral classifiers represent a unique and distinguished set of resilience-linked phenotypes.

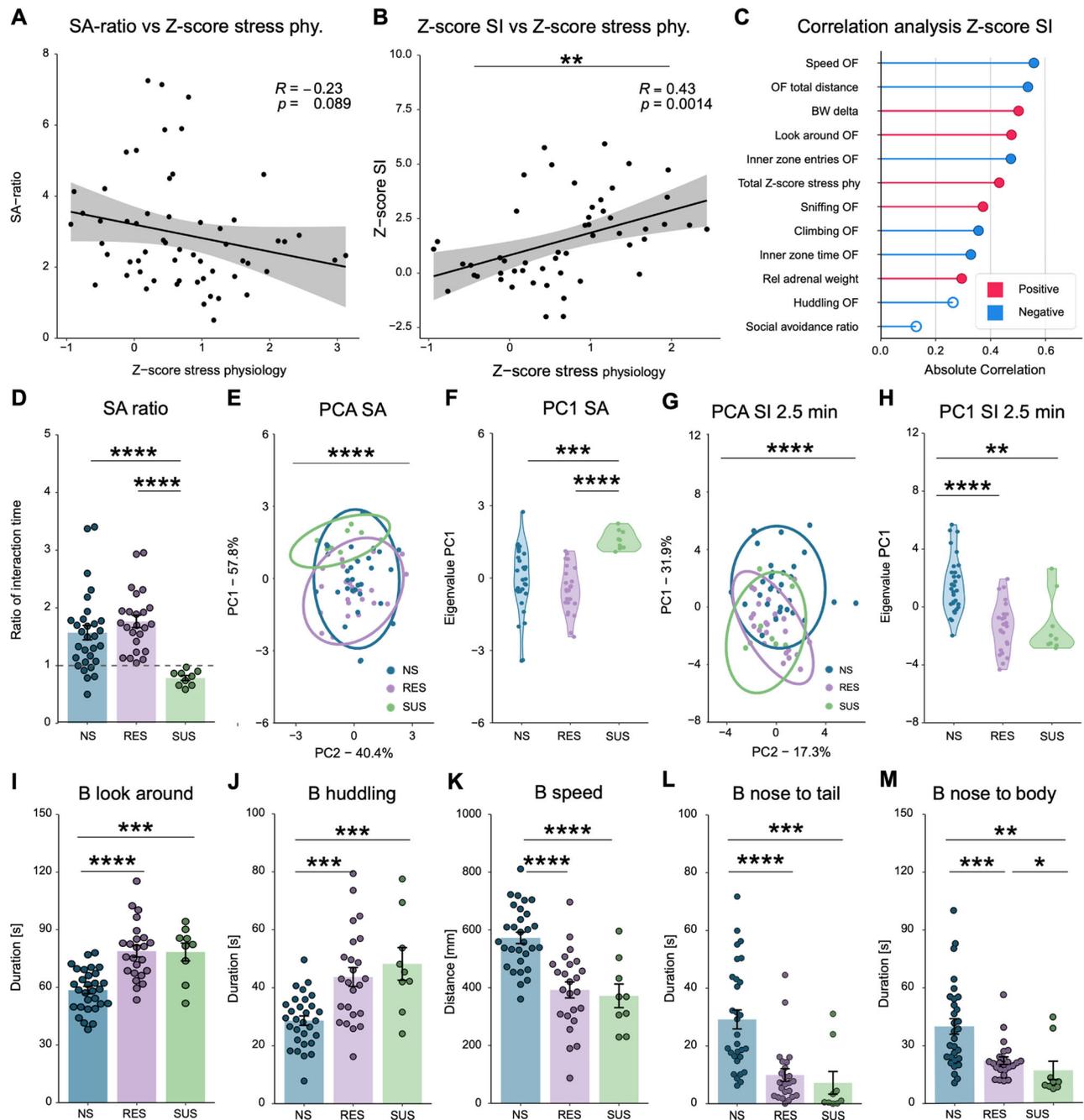
The DeepOF unsupervised pipeline can be flexibly applied across different experimental settings

The unsupervised pipeline within DeepOF was applied to three datasets and four settings. These included both single and multi-animal embeddings on the SI dataset, single-animal embeddings on the OF dataset, and single-animal embeddings on the SA dataset. When applying this workflow to a new dataset, the number of clusters is a hyperparameter the user must tune. In this study, an optimal solution was found by selecting the number of clusters that explains the largest difference between experimental conditions (in terms of the area under the ROC curve of a classifier to distinguish between them, see methods for details). While DeepOF could be used to describe the behavioral space of a single condition, this model selection procedure aims at maximizing the power to detect behavioral differences between experimental conditions. An optimum of 10 clusters was measured for both single- and multi-animal SI settings (Fig. 6A and

Supplementary Fig. 5A), whereas the single-animal OF setting showed an optimum of 11 clusters (Supplementary Fig. 6A), and the SA setting of 17 clusters (Supplementary Fig. 7A). Timepoint UMAP projections of the latent space depicting all clusters can be found in Fig. 6B, and Supplementary Figs. 5B, 6B, and 7B for all four settings, respectively.

DeepOF can quantify behavioral differences over time in an unsupervised way

Once the number of clusters was fixed, the stress-induced phenotype was investigated over time in both SI and OF settings. SA was excluded of this analysis due to the shorter length of the videos (2.5 min), in which no decay of arousal should be observed in the animals. To this end, a growing time window spanning an increasing number of sequential seconds was analyzed. For each analysis, the discriminability between conditions was tested by evaluating the performance of a linear classifier to distinguish between them in the global animal embedding space, for which each experiment is represented by a vector containing the time spent per cluster (see methods for details). The bin size for which discriminability was maximized was then selected as optimal and used for further analysis. In this case, we observed an optimum of 126 and 124 s for the single-animal and multi-animal SI tasks respectively, indicating that differences between conditions are maximized early in the 10-min-long experiments, which is compatible with habituation. Furthermore, performance across



consecutive, non-overlapping bins retaining the optimal size was also reported (Fig. 6C and Supplementary Fig. 5C). Here, decaying performance across bins in the SI setting is also compatible with a state of arousal, where conditions become less distinguishable over time after the behavior of the C57Bl/6N mice becomes less influenced by novelty. The largest difference between NS and CSDS animals can thus be observed during this period. In line with this finding, the optimal distance in the single animal OF data was reached at 595 s, suggesting that no binning is necessary since behavior between conditions remains consistently distinguishable across the videos (Supplementary Fig. 6C).

Interestingly, global animal embeddings show a clearer separation between conditions in both single and multi-animal embeddings for the SI setting (Fig. 6D and Supplementary Fig. 5D), whereas the difference is milder in the OF setting, as the projected distributions are less separable (Supplementary Fig. 6D). In the SA setting, projections

show, as expected, a higher separation between conditions in trial two, which includes the engaged conspecific (Supplementary Fig. 7C, D).

These global embeddings also capture how distributions merge over time in the SI settings, as the behavioral profiles of NS and CSDS mice become closer (Supplementary Figs. 8, 9).

Individual unsupervised clusters reveal differences in behavior enrichment

Going beyond global differences in behavior, the aggregated embeddings depicted so far are the result of summarizing the expression of the set of detected behavioral clusters. Once obtained, DeepOF enables the user to test the differential expression between conditions. To this end, the time spent on each cluster across all videos for each condition is recorded for each time bin. Importantly, DeepOF has no knowledge of the assigned animal conditions at the time of training and assigning clusters.

Fig. 5 | Z-score correlation analysis and the exploration of susceptibility and resiliency. **A** Pearson correlation analysis between the SA-ratio and the Z-score of stress physiology ($R = -0.23, p = 0.089$). **B** Pearson correlation analysis between the SI task 2.5 min time bin top five contributing behaviors and the Z-score of stress physiology ($R = 0.43, p = 0.0014$). **C** Pearson correlation analyses between the Z-score of SI and all other parameters. A strong correlation was observed with several OF parameters, such as speed ($R = -0.56, p = 1.76e-5$), total distance ($R = -0.54, p = 4.27e-5$), look-around ($R = 0.48, p = 0.0004$), and inner zone: entries ($R = -0.47, p = 0.0004$), but not with the SA-ratio ($R = -0.13, p = 0.37$). **D** The SA-ratio shows a significant main effect with the Kruskal-Wallis: $H(2) = 21.22, p < 0.0001$. Wilcoxon posthoc shows that SUS animals (SI-ratio < 1) have a significantly lower SI-ratio compared to NS animals $W(9,30) = 249, p = 4.1e-5$ and RES animals $W(9,24) = 216, p = 1.56e-7$. There is no difference between NS and RES animals $W(30,24) = 270, p = 0.12$. **E** The PCA for SA shows a significant main effect with the one-way ANOVA: $F(2,60) = 10.90, p = 9.19e-5$. **F** The PCI eigenvalues of the SA show a significant difference between SUS and NS animals Post-hoc Benjamini Hochberg (BH): $T(9,30) = p = 0.0005$ and between SUS and RES animals $T(9,24) = p = 5.88e-5$. There is no significant difference between NS and RES animals $T(30,24) = p = 0.196$. **G** The PCA for the 2.5 min SI ratio shows a significant main effect with the Kruskal-Wallis: $H(2) = 24.83, p = 4.06e-6$. **H** The PCI eigenvalues of the 2.5 min bin SI show a significant difference between NS and RES animals Post-hoc Wilcoxon: $W(30,24) = 92, p = 1.82e-6$, and between NS and SUS animals $W(30,9) = 41, p = 0.0015$. There is no difference between RES and SUS animals ($W(24,9) = 117, p = 0.736$). **I** B-look-around shows a significant main effect with the one-way-ANOVA: $F(2,60) = 19.23, p = 3.53e-7$. Post hoc BH shows a significant difference between NS and RES ($T(30,24) = p = 9.86e-7$), and NS and SUS ($T(30,9) = p = 0.0002$), but no difference between RES and SUS ($T(24,9) = p = 0.94$).

J B-huddle shows a significant main effect with the one-way-ANOVA: $F(2,60) = 12.35, p = 3.23e-5$. Post hoc BH shows a significant difference between NS and RES ($T(30,24) = p = 0.0003$), and NS and SUS ($T(30,9) = p = 0.0004$), but no difference between RES and SUS ($T(24,9) = p = 0.39$). **K** B-speed shows a significant main effect with the one-way-ANOVA: $F(2,60) = 18.63, p = 5.1e-7$. Post hoc BH shows a significant difference between NS and RES ($T(30,24) = p = 3.12e-6$), and NS and SUS ($T(30,9) = p = 7.62e-5$), but no difference between RES and SUS ($T(24,9) = p = 0.67$). **L** B-nose-to-tail shows a significant main effect with the Kruskal-Wallis: $H(2) = 26.70, p = 1.59e-6$. Post hoc Wilcoxon shows a significant difference between NS and RES ($W(30,24) = 628, p = 1.82e-6$), and NS and SUS ($W(30,9) = 236, p = 0.0005$), but no difference between RES and SUS ($W(24,9) = 152.5, p = 0.075$). **M** B-nose-to-body shows a significant main effect with the Kruskal-Wallis: $H(2) = 19.61, p = 5.52e-5$. Post hoc Wilcoxon analysis shows a significant difference between NS and RES ($W(30,24) = 567, p = 0.0003$), and NS and SUS ($W(30,9) = 230, p = 0.0009$), and RES and SUS ($W(24,9) = 167, p = 0.018$). The correlation analyses (A, B) are represented with a regression line and a 95% confidence interval window and all individual samples as points. **C** has the correlation value (R) represented by the red line (positive) or blue line (negative), black circles around the points are identified as significant correlations, $p < 0.05$. The bar graphs are presented as mean \pm standard error of the mean and all individual samples as points. The PCA graphs (E, G) are provided with a 95% confidence ellipse and all individual samples as points. Further PCI analyses are represented with a violin plot and all individual samples as points (F, H). The bar graphs are presented as mean \pm standard error of the mean and all individual samples as points. $N = 30$ for NS and CSDS in (A), and $n = 26$ for NS and $n = 27$ for CSDS in (B, C), $n = 30$ for NS, $n = 24$ for RES, $n = 9$ for SUS in (D–M). Source data are provided as a Source Data file.

The expression between NS and CSDS animals was then compared using 2-way Mann-Whitney U tests for each cluster independently, and p values were corrected for multiple testing using the BH method across both clusters and time bins, when applicable. We observed significant differences in eight out of ten and six out of ten clusters for the first time bin of the single and multi-animal SI settings, respectively (Fig. 6E and Supplementary Fig. 5E). Interestingly, and in line with habituation to the environment, these differences also fade across time. The single-animal setting still shows some (although less) significant differences in all time bins, albeit with reduced effect sizes (Supplementary Fig. 10). Interestingly, also in the single-animal embeddings, cluster 8 remains highly significant during the entire course of the experiments. The multi-animal setting yields in contrast almost no significant results beyond the first time bin (Supplementary Fig. 11).

In the OF setting, 7 out of 11 clusters showed a significant differential expression in the first 595 s (Supplementary Fig. 6E). The SA test, in turn, is an interesting setting to test DeepOF given that its main axis of variation is the distance to the cage with the conspecific, which constitutes information that is not available to DeepOF in its current form (which only looks at the posture of the tracked animals). Interestingly, and while the analysis shows no significant results in trial one (without the conspecific, Supplementary Fig. 7E), 6 out of 17 clusters show significant differential expression in trial two (with the conspecific, Supplementary Fig. 7F), suggesting that DeepOF can correctly detect behavioral differences even without absolute location information.

Finally, we also explored the spatial distribution of cluster expression across all three settings. We obtained heatmaps depicting the global exploration of the arena by the C57Bl/6N across all videos (for both conditions). Along these lines, our results show how, while, as shown, CSDS animals tend to occupy the center of the arena significantly less (Fig. 2F) there is no spatial preference across animals for individual clusters (Fig. 6F and Supplementary Figs. 5F, 6F show the overall locomotion distribution, while a comprehensive overview of individual clusters is presented in Supplementary Figs. 12, 13, and 14).

Individual unsupervised clusters reveal differences in behavior dynamics

Aside from comparing cluster enrichment, DeepOF can help gain insight into how cluster transitions and sequences differ across conditions. To accomplish this, an empirical transition matrix was obtained for each condition by counting how many times an animal goes from one given cluster to another (including itself). Since all transitions were observed to have non-zero probability, the Markov chains obtained from simulations can be proven to reach a steady state over time (where probabilities to go from one behavior to another stabilize). The entropy of these steady state distributions was reported for both conditions, with higher values corresponding to a less predictable exploration of the behavioral space. Interestingly, CSDS animals showed a significantly lower behavioral entropy in the social interaction task than their NS counterparts, retrievable in both single and multi-animal embeddings (Fig. 6F and Supplementary Fig. 5F). This goes in line with the NS animals exploring the behavioral space more thoroughly, while CSDS animals are more conditioned by the conspecific. In line with this hypothesis, no significant differences across conditions were found in the single-animal OF experiments (Supplementary Fig. 6F). Moreover, to validate these results, the obtained behavioral entropy score was correlated with the physiology Z-score presented earlier (Supplementary Fig. 15). As expected, significant negative correlations were found for the SI setting both when exploring the single and multi-animal behavioral spaces. No significant correlation was observed for the single-animal OF setting.

Shapley additive explanations reveal a consistent profile across differentially expressed clusters

An important aspect of any machine learning pipeline using highly complex models is its explainability. In this study, we aimed to explain cluster assignments by fitting a multi-output supervised classifier (a gradient boosting machine) that maps statistics of the initial time series segments (including locomotion and individual body part areas, speeds, distances, and angles) to the subsequent cluster assignments. Performance and generalizability of the constructed classifiers across the dataset were assessed in terms of the balanced accuracy on a 10-

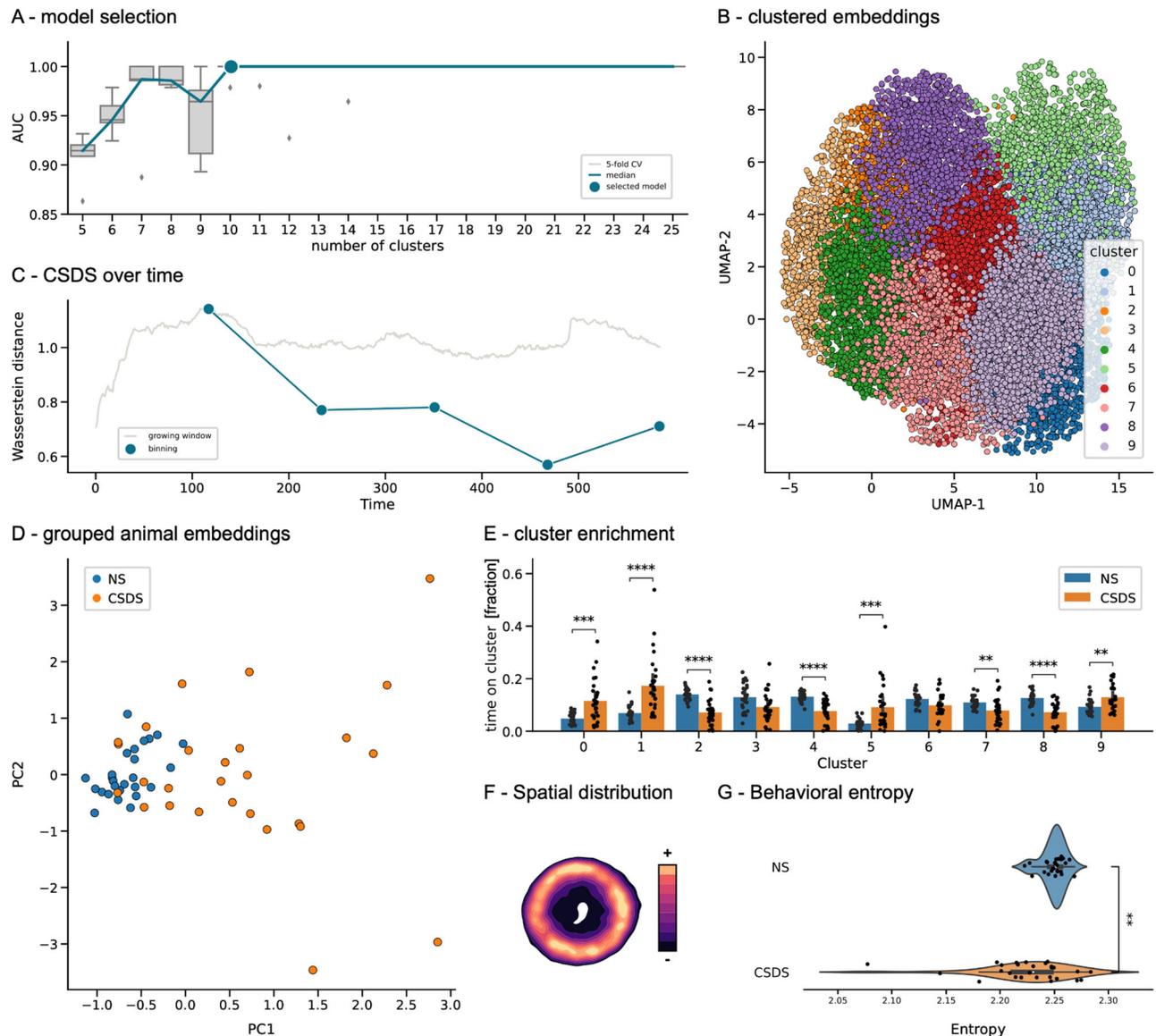


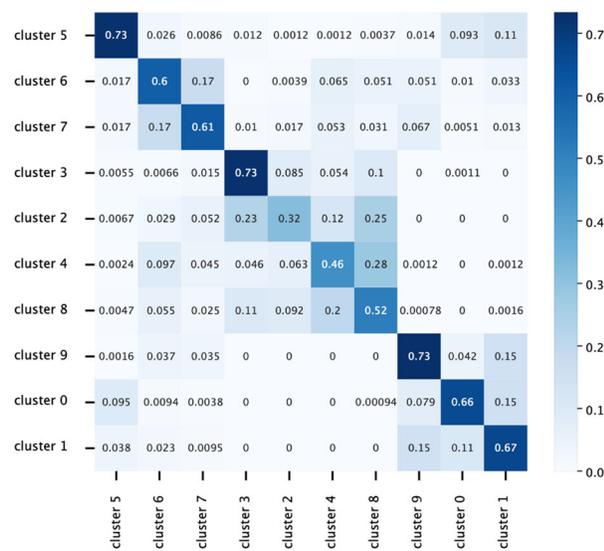
Fig. 6 | Single-animal unsupervised analyses identify different behavioral patterns between stressed and non-stressed mice during the SI task. **A** Cluster selection pipeline results reporting the area under the ROC curve from a logistic regression classifier discriminating between conditions. A 10-component solution (from a range between 5 and 25) was selected as optimal in a fivefold ($N = 5$) cross-validation loop (see methods for details). **B** Embeddings by time point obtained using DeepOF’s unsupervised pipeline. Different colors correspond to different clusters. Dimensionality was further reduced from the original 8-dimensional embeddings using UMAP for visualization purposes. **C** Optimal binning of the videos was obtained as the Wasserstein distance between the global animal embeddings of both conditions across a growing window, between the first 10–600 s for each video at one-second intervals (gray curve). Higher values correspond to larger behavioral differences across conditions. A maximum was observed at 126 s, close to the stipulated 150 s selected based on the SA task literature. The dark green curve depicts the Wasserstein distance across all subsequent non-overlapping bins with optimal length. The decay observed across time is consistent with the hypothesized arousal period in the CSDS cohort. **D** Representation of the global animal embeddings for the optimally discriminant bin (126 s) per experimental video colored by condition (see methods for details).

E Cluster enrichment per experimental condition ($N = 26$ for NS and $N = 27$ for CSDS) in the first optimal bin (first 126 s). Reported statistics correspond to a 2-way Mann-Whitney U non-parametric test corrected for multiple testing using Benjamini-Hochberg’s method across both clusters and bins (significant differences observed in clusters 0: $U = 1.6e+2$, $p = 7.7e-4$, 1: $U = 1.1e+2$, $p = 1.3e-5$, 2: $U = 6.3e+2$, $p = 1.1e-6$, 4: $U = 6.4e+2$, $p = 3.3e-7$, 5: $U = 1.6e+2$, $p = 6.3e-4$, 7: $U = 5.3e+2$, $p = 1.3e-3$, 8: $U = 6.2e+2$, $p = 1.9e-6$, 9: $U = 1.9e+2$, $p = 4.4e-3$). Bar graphs represent mean \pm standard deviation of the time proportion spent on each cluster. **F** Example heatmap depicting spatial distribution across all experiments (in both conditions) for all clusters. Specific heatmaps for all individual clusters are available in Supplementary Fig. 12). **G** Behavioral entropy scores per condition. NS animals show a significantly higher entropy than CSDS animals, which can be attributed to a less predictable exploration of the behavioral space ($U = 5.3e+2$, $p = 1.68e-3$, $N = 26$ for NS and $N = 27$ for CSDS). Moreover, and in accordance with these results, behavioral entropy shows a significant negative correlation with the presented stress physiology Z-score (Supplementary Fig. 15A). Source data are provided as a Source Data file. Box plots in (A, G) show the median and the inter-quartile range. Whiskers show the full range, excluding outliers as a function of the inter-quartile range.

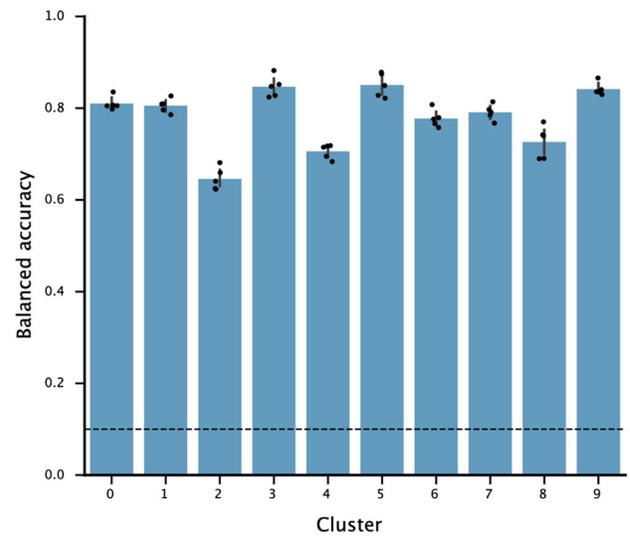
fold stratified cross-validation loop, which was designed so that segments coming from the same video were never assigned to both train and test folds. Data for SI (single and multi-animal) and OF settings were standardized, and the minority class was oversampled using the

SMOTE algorithm to correct for class imbalance. Performance per cluster is shown by means of the confusion matrices per task and the balanced accuracy per cluster (Fig. 7A, B and Supplementary Figs. 16A, B and 17A, B for all three settings, respectively). Importantly, classifier

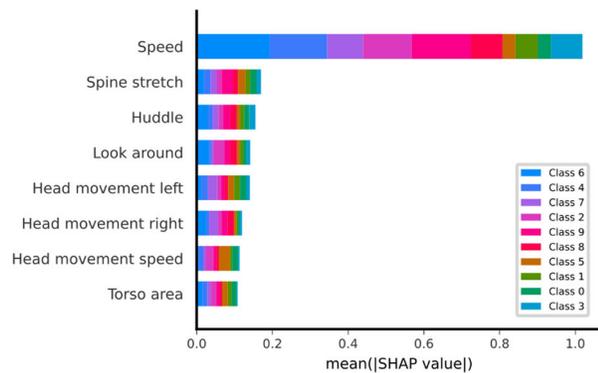
A - cluster detection confusion matrix



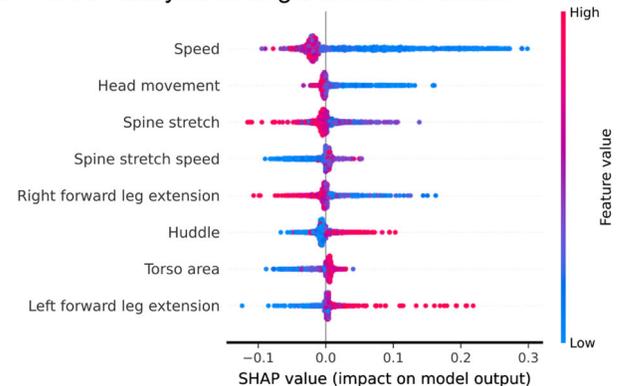
B - cluster detection performance



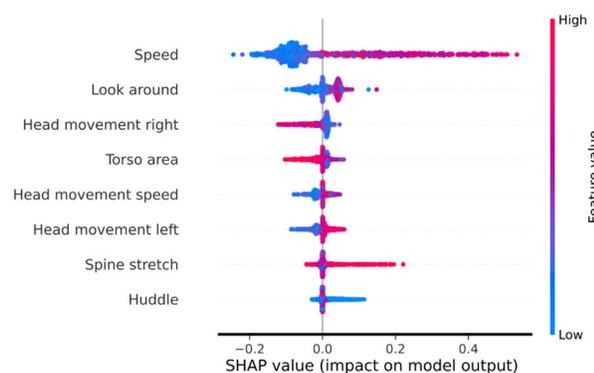
C - SHAP global feature importance



D - SHAP analysis of single-animal SI cluster 1



E - SHAP analysis of single animal SI cluster 2



F - SHAP analysis of single animal SI cluster 8

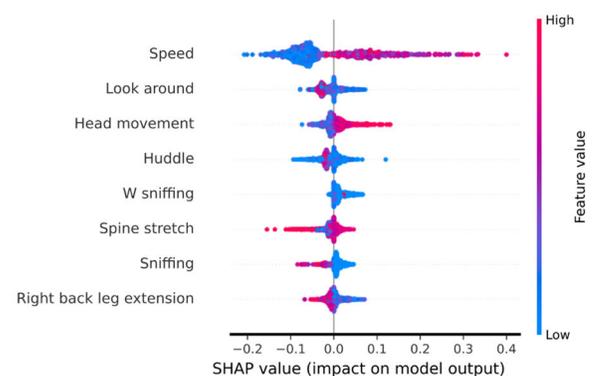


Fig. 7 | SHAP analysis of unsupervised cluster assignments in the single-animal social interaction task. Gradient boosting machines were trained to map from a predefined set of time series statistics (including body part speeds, distances, distance speeds, areas, area speeds, and supervised annotations) to the previously obtained cluster assignments. **A** Confusion matrix obtained from the trained gradient boosting machine classifying between clusters. Aggregated performance over the validation folds of a fivefold cross-validation is shown. **B** Validation performance per cluster across a fivefold ($N = 5$) cross-validation loop. Balanced accuracy was used to correct for cluster assignment imbalance. The dashed line marks the expected performance due to chance, considering all outputs. Bars show

mean \pm 95% confidence interval. **C** Overall feature importance for the multi-output classifier using SHAP. Features in the y-axis are sorted by overall absolute SHAP values across clusters. Classes on the bars are sorted by overall absolute SHAP values across features. **D–F** Bee swarm plots for the three most differentially expressed clusters between NS and CSDS mice (1, 2, and 5), identified with the unsupervised DeepOF pipeline on the SI experiments using single-animal embeddings. The depicted plots display the first eight most important features for each classifier, in terms of the mean absolute value of the SHAP values. Source data are provided as a Source Data file.

performance is substantially greater than random in all cases for all three settings, meaning that all clusters are highly distinguishable from one another by the set of summary statistics employed.

The result of this analysis is thus a set of feature explainers for each retrieved cluster, which can be used to interpret, alongside visual inspection of the corresponding video fragments (included as Supplementary files), what the obtained behavioral motifs represent. Both global (Fig. 7C, Supplementary Figs. 16C, 17C) and cluster-specific feature importance values can be retrieved. In this context, we found consistent descriptions of clusters that are differentially represented across conditions for all three tasks.

In the single-animal SI task, for example, cluster 1 (Fig. 7D, enriched in CSDS animals) is consistently explained by low locomotion speed, low head movement, and low spine stretch, and is positively associated with the huddle classifier. Visual inspection reveals a behavior close to freezing. Cluster 2 (Fig. 7E, enriched in NS animals) is in contrast explained by high locomotion speed, exploratory behavior, low head movement, and spine stretch. Close visual inspection depicts active locomotion and engagement with the conspecific. Interestingly, cluster 8 (Fig. 7F, enriched in NS animals across all time bins) is explained by increased speed, head movement, and negatively associated with sniffing. Visual inspection suggests engaging in motion (shifting from a still position to active locomotion).

In the case of the multi-animal SI setting, the explainability pipeline reveals how the models work differently when taking both animals into account. In this case, the two-animal system is embedded as a whole, and features including both animals are considered when running SHAP. As mentioned in the methods section, a regularization hyperparameter allows the system to focus more on interactions between the animals or in joint individual behaviors. In this case, we used a moderated value of the parameter that enables the contribution of both, which becomes apparent when analyzing the explainability profiles of the retrieved behaviors. Cluster 3, for example (Supplementary Fig. 16D, highly enriched in CSDS), is explained not only by low speed on the C57Bl/6 N animal, but also by increased speed of the CD1, among others. Upon visual inspection, one can observe exactly that the CD1 is exploring the arena while the C57Bl/6N stands still, in a posture usually associated with the stopped and huddled trait. Cluster 5 (Supplementary Fig. 16E, also enriched in CSDS) closely captures an interaction between the two animals, where the CD1 is typically more engaged in movement. The SHAP pipeline eloquently reveals negative correlations with spine stretch and back, torso, body and head areas, as well as speed in both mice. Conversely, cluster 8 (Supplementary Fig. 16F, enriched in NS) is well explained by increased speed in both animals, which can be confirmed by visual inspection.

Finally, this pipeline was also used to interpret clusters in the OF setting. In this case, cluster 0 (Supplementary Fig. 17D, enriched in CSDS animals) is explained by a decreased overall speed, positive correlations with mid and back spine stretch, back area, and left leg extensions, and negative association with right leg extensions. Visual inspection indeed reveals a cluster highly enriched in digging. Cluster 8 (Supplementary Fig. 17E, also enriched in CSDS animals), is in turn explained by decreased speed, mid, and back spine stretch, increased head area and extended right legs. Visual inspection shows a cluster enriched in slow walking, often including head movement and interaction with the walls. Finally, cluster 9 (Supplementary Fig. 17F, enriched in NS animals) is positively correlated with speed and head movement, and negatively correlated with spine stretch, among others. Visual inspection depicts an exploratory behavior with active movement.

All in all, the provided cluster explainability pipeline is a useful tool to interpret all reported patterns. Moreover, visual inspection of cluster snippets is also made possible with a single command within DeepOF, which makes the interpretation process more effective.

Discussion

For decades there has been a trend to standardize and simplify social behavioral tests, which has led to an oversimplification of the description of the social behavioral repertoire. The current developments of open-source markerless pose estimation tools for tracking multiple animals have provided the possibility for more complex and socially relevant behavioral tests. The current study provides an open-source tool, DeepOF, which can investigate both the individual and social behavioral profiles in mice using DeepLabCut-annotated pose estimation data. Applying this tool, the current study identified a distinct social behavioral profile following CSDS using a selection of five traits annotated by DeepOF on the C57Bl/6N animal. In addition, a similar social behavioral profile was identified using an unsupervised workflow, which could detect behavioral differences in different experimental settings, including social interaction and single-animal open field tests, and a social avoidance task. Moreover, DeepOF allowed to study behavioral dynamics in unprecedented detail and identified the 5 min during the interaction with a novel conspecific as crucial for the social profiling of CSDS exposure in both supervised and unsupervised workflows. Overall, this study demonstrates the high utility and versatility of DeepOF for the analysis of complex individual and social behavior in rodents.

DeepOF as part of a markerless pose estimation toolset

The initial release of DeepLabCut in 2018²⁹ provided a reliable and accessible tool for researchers around the globe to process markerless pose estimation data, which has undoubtedly changed the field of behavioral neuroscience. This has set in motion a rapid growth of tools for analyzing pose estimation data that are increasing the range of possibilities in the field, which were unimaginable using classical tracking approaches or manual scoring. An important distinction between these pose estimation analysis tools is whether they intend to extract pre-defined and characterized traits (supervised) or to explore the data and extract patterns without external information (unsupervised). The DeepOF module is designed to provide both analysis pipelines. The supervised behavioral classifiers offer a quick and easy-to-use analysis to detect individual and social behavioral traits without manual labeling. In addition, when differences between the conditions are not reflected in these traits, or the researcher aims to obtain behavioral embeddings, the DeepOF package can encode the data in a time-aware way that can report differentially expressed patterns in an unsupervised manner, taking single and multi-animal inputs.

The supervised framework: spotting recognizable patterns

The supervised pipeline within the DeepOF package can be used on single and dyadic behavioral data in multiple-shaped arenas. DeepOF is capable of reporting a pre-defined set of behavioral traits without any extra labeling or training required. To accomplish this, it relies on both simple rule-based annotations and machine learning binary classifiers whose generalizability has been tested, trading off flexibility for ease of use. This makes it user-friendly for researchers without computational expertise to apply this supervised pipeline, without having to make any modifications. To further detect unsupported patterns, using a more involved and flexible tool (such as SimBA³⁷ or MARS²⁷) could be a reasonable next step to take. These tools include a supervised approach that requires the user to label and train classifiers, providing the freedom to train powerful classifiers and recognize behavioral traits, which is especially beneficial for labs without computational expertise. However, in contrast to DeepOF, this approach also delegates to the user the responsibility of testing the generalizability of the results (how well the trained models can be applied to newly generated data, even in similar settings), which requires careful practices from the experimenters.

The DeepOF module provides a more complete social behavioral profile than the social avoidance task

The social behavioral profile in CSDS-subjected animals has been measured extensively using the SA task, which is based on the separation of social behavioral traits between non-stressed and stressed animals^{11,17,38}. Previous research has shown that rodents have a social interaction preference towards a novel conspecific compared to a familiar conspecific³⁹. However, the duration of this social behavioral arousal state has not been well documented. In this context, and by replicating the time the SA task typically lasts for¹⁰, the current study shows that the CSDS-related social behavioral profile, obtained with the DeepOF supervised classifiers, was increasingly observed during the first 2.5 min of the 10 min SI task. Furthermore, the presented unsupervised workflow was used to determine an optimal binning of our experiments by measuring how different both conditions were across time for a linear classifier. This yielded an optimal separation at -2.1 min (126 and 124 s when testing with single and multi-animal embeddings, respectively), which then decayed over subsequent time bins in a manner consistent with the arousal hypothesis. The fact that this result was not seen in the absence of a conspecific strengthens this argument. Taking this into account, we argue that the introduction of a novel conspecific induces a state of arousal, which coincides with a distinct social behavioral profile that disappears over time after 2–3 min due to habituation.

Along these lines, this study shows that the DeepOF social behavioral classifiers provide a stronger separation of the social behavioral profile between stressed and non-stressed animals compared to the classical SA task, which also correlates better to physiological stress parameters.

Furthermore, the identification of stress-susceptible and resilient animals is often performed using the SA-ratio of the SA task^{10,17} and for this DeepOF offers unique advantages. While the SA ratio clearly distinguishes stress-affected individuals, especially following more severe CSDS paradigms, the DeepOF module will significantly advance the possibilities and sensitivity of this distinction, by investigating the degree of resilience based on multiple behavioral classifiers with high sensitivity and in freely moving animals, which enables uncovering a so far undescribed set of resilience-linked phenotypes that are different from the univariate SA task. Taken together, it can be concluded that using the DeepOF social behavioral classifiers provides a more robust and clearer social behavioral profile in animals subjected to CSDS compared to the SA task. An important reason for the superiority of DeepOF in social behavioral profiling depends on the experimental setup: the SA task relies on the confinement of an animal (for example using a wired mesh cage), which means that no natural interaction between freely moving animals is possible, whereas the SI task is based on a naturally occurring interaction between freely moving animals¹⁸. Moreover, in the SA task, the confined animal can show symptoms of anxiety-related behavior, which influences the physiological state and the social interaction and approach behavior of the conspecific^{40–42}. Differences in anxiety-related behavior between experimental animals can still contribute to alterations in social behavior and recent data suggest distinct neurobiological circuits driving both phenotypes⁴³, therefore sufficient habituation and the ability to observe behavior in freely moving animals will lead to improved discrimination. Moreover, a further crucial advantage of the DeepOF module is the many different behavioral classifiers that can be investigated at the same time without increasing the labor intensity. The combined analysis of multiple behavioral classifiers into a Z-score of social behavior provides a more complete social behavioral profile than solely investigating social avoidance behavior.

DeepOF can detect and explain differences across experimental conditions in a fully unsupervised way, embedding data from one or more animals

The supervised pipeline within DeepOF follows a highly opinionated philosophy, which focuses on ease of use and relies on predefined

models. As an alternative, DeepOF offers an unsupervised workflow capable of representing animal behavior across experiments without any label information. In its most basic expression, this involves obtaining a representation for each experiment in a time-aware manner: unlike other dimensionality reduction algorithms like PCA, UMAP, and T-SNE²⁶, DeepOF, when applied to the raw dataset, relies on a combination of convolutional and recurrent neural networks capable of modeling the sequential nature of motion. Each input to the models consists of a subsequence across a non-overlapping sliding window of each experiment. Although this idea has been explored before³³, DeepOF introduces several novelties to the field, such as unified embedding and clustering, the support for multi-animal embeddings, and graph representations that integrate not only coordinates by also body-part-specific speed and distance information, which makes it ideal for settings where informative body parts (such as paws) are occluded, as is the case for commonly used top-down videos.

In addition, these global embeddings can be decomposed into a set of clusters representing behavioral motifs that the user can then inspect both visually and with machine learning explainability methods. Moreover, by comparing cluster enrichment and dynamics across conditions, it is possible to answer questions that are relevant to understanding what the observed difference might be based on, without any previous knowledge: Which behaviors are most or least expressed in each condition? Is the set of behaviors expressed differently in experimental conditions? Are they expressed differentially across space and time? This constitutes a complementary approach that can be beneficial to further direct hypotheses when little knowledge is available. In addition, by not only showing overall differences between cohorts but also reporting which motion primitives might be driving them, it is possible to test hypotheses by training novel supervised classifiers based on those motion primitives. This can allow researchers to distinguish new, meaningful patterns that have not been reported before and that may be significantly associated with a given condition.

Taken together, the current study exemplifies that the unsupervised pipeline provided in DeepOF does not only recapitulate results previously obtained with the supervised analysis, but also shows how this tool can be used to detect habituation and overall differences in behavioral exploration. We also show that detected differences are significantly stronger when a conspecific is present, although also detectable during single animal arena exploration alone.

Towards an open-source behavioral analysis ecosystem

One of the main advantages of DeepOF, SimBA³⁷, VAME³³, MARS²⁷, and many other packages cited in this manuscript, is that they are open source. This means that their inner workings are transparent, and that it is possible for the community to contribute to their development. We strongly believe that the adoption of open-source frameworks can not only increase transparency in the field but also incentivize a feeling of community, in which researchers and developers can share ideas, code, and solve bugs and problems together. Moreover, the open source framework facilitates beneficial feedback loops, where the data generated using these tools can be published, thus increasing the opportunity to produce better software. A good example of this is zero-shot pose estimation⁴⁴, which enables motion tracking without labeling, by cleverly leveraging information from several publicly available datasets. In addition, new technologies are starting to enable joint learning from multiple modalities, such as neural activation and behavior⁴⁵, which enables the exploration of how these modalities are influencing each other.

In addition to the software, an equally important problem to tackle is the need for open-source benchmarks. As platforms for testing and validating pose estimation and detection algorithms become available, it becomes easier to clearly show and compare the performance of different software options for different tasks. An

example of this is the Caltech Mouse Social Interactions (CalMS21) dataset, a pioneer in the field that provides benchmarking for classic detection of social interactions, annotation style transfer, and detection of rare traits⁴⁶. While unsupervised learning benchmarking remains highly unexplored to the best of our knowledge, it would be crucial to compare the DeepOF pipeline with other available methods in this context when the tools become available.

Finally, and in contrast to several other options that offer extended functionality but rely on proprietary algorithms and/or specialized hardware²³, these tools have the potential to make otherwise expensive software available to a larger audience.

In conclusion, the current study provides a novel approach for individual and social behavioral profiling in rodents by extracting predefined behavioral classifiers and unsupervised, time-aware embeddings using DeepOF. Furthermore, while the tool provides means of customization, it is uniquely optimized for the most common behavioral setup: top-down video recordings. Moreover, we show evidence for the validation of the provided behavioral annotators and offer an open-source package to increase transparency and contribute to the further standardization of the behavioral constructs. We also show that, while differences across conditions are detectable during single animal exploration, they are enhanced in the SI task involving a companion mouse. Furthermore, while the classical SA task does identify the social behavioral profile induced by CSDS, the DeepOF behavioral classifiers provide a more robust and clearer profile. DeepOF is thereby a highly versatile tool that can also be applied to other research questions, e.g., to study sex differences in social behavior or analyze home-cage behavior throughout the lifespan of animals using longitudinal recordings. In addition, the DeepOF module contributes to a more specific classification of the affected individual and social behaviors in stress-related disorders, which could contribute to the study of drug development for psychiatric disorders.

Methods

Time series extraction from raw videos

Time series were extracted from videos using DeepLabCut version 2.2b7 (single animal mode). 11 body parts per animal were tagged, including the nose, left and right ears, three points along the spine (including the center of the animal), all four extremities, and the tail base (Fig. 1A). The DeepLabCut model was trained to track up to two animals at once (one CD1 mouse and one C57Bl/6 N mouse) and can be found in the Supplementary material (see code and data availability statement). Using the multi-animal DeepLabCut³⁰, extending the tracking to animals from the same strain is also possible. Next, DeepLabCut annotated datasets were processed and analyzed using DeepOF v0.4.6³⁶.

Time series data preprocessing

All videos and extracted time series undergo an automatic preprocessing pipeline that is included within the DeepOF package, consisting of smoothing and two sequential imputation levels, applied to all body parts of all tracked animals independently. For smoothing DeepOF applies a Savitzky-Golay filter⁴⁷ to each independent tracked variable by fitting an $n/2$ -degree polynomial over an n -frame sliding window, where n is the frame rate of the corresponding videos.

To identify and correct any artifacts in the time series, a moving average model is then fitted to the time-based quality scores of each tracked variable (as reported by DeepLabCut's output likelihood). By detecting divergences (of at least three standard deviations) from the moving average model, DeepOF can detect sudden and consistent drops in tracking quality, often correlated with body-part occlusions. Body parts with low quality are thus removed from the data, and further imputed using sci-kit learn's iterative imputer with default parameters⁴⁸, which predicts missing values based on all available

features at a given time point using a Bayesian ridge regression method. A second imputation method is then conducted, aiming to remove spatial jumps in the tracked body parts. To do this, another moving average model is fitted, this time to the body part coordinates themselves, and any data point located at least three standard deviations from the model is replaced by the predicted values.

Time series feature extraction

After preprocessing the time series independently, DeepOF extracts a set of features aiming to describe how entire animals move and interact. These include centered and aligned coordinates, distances between body parts, angles, and areas of specific regions of each available body (Fig. 1B), as well as their speeds, accelerations, and higher-order derivatives. The value for each feature is reported per time point.

Coordinates. Raw coordinates for each body part are centered (the cartesian origin is set to the center of each animal) and vertically aligned so that the y -axis matches with the line delimited by the *center* of each animal and *spine 1* (see Fig. 1A for reference). This is done so that both translational and rotational variances are not considered in further processing steps (in principle, and except for some annotations such as wall climbing and sniffing—see below—DeepOF extracts posture patterns that are invariant to where in the arena and in which rotational orientation they are expressed).

Distances and angles. Distances and angles over time between all body parts within and across all animals are computed by DeepOF by default, and available for retrieval.

Areas. The full area of the animal over time is computed by DeepOF by defining a polygon on all external body parts (*nose*, *ears*, *legs*, and *tail base*). The head area is delimited by the *nose*, *ears*, and *spine 1*. The Torso area is delimited by *spine 1*, both *forward legs*, and *spine 2*. The back area is delimited by the *center*, both back legs, and the *tail base*.

Finally, speeds, accelerations, jerks, and larger-order derivatives of each extracted feature are also computed using a sliding window approach. Importantly, the detailed 11-body-part labeling scheme suggested and provided by DeepOF plays a crucial role here. While parts of the pipeline can still work with fewer labels, the comprehensive set of features that DeepOF is able to extract with this set of labels enhances not only supervised annotations, but also data representations and model interpretability.

Supervised behavioral tagging with DeepOF

The supervised pipeline within DeepOF aims to provide a set of annotators that work out of the box (without user labeling) for several behaviorally relevant traits. The workflow supports both dyadic interactions and individual traits, which are reported for each mouse individually (Fig. 1C). Furthermore, annotated traits fall into one of two categories:

1. *Traits annotated based on pre-defined rules.* Several motifs of interest are annotated using a set of rules that do not require a trained model. For example, contact between animals can be reported when the distance between the involved body parts is less than a certain threshold.
2. *Traits annotated following a supervised machine learning pipeline.* While rule-based annotation is enough for some traits, others are too complex or might be manifested in subtly different ways, and machine learning models are often a better option. In this case, a rigorous validation pipeline has been applied to measure the performance of the classifier not only in a separate test data set, but also across datasets comprehending different arenas and laboratories.

Rule-based annotated traits. Among the rule-based annotated dyadic traits, nose-to-nose and nose-to-tail depend on single distance thresholds between specific body parts of the animals involved. In the case of nose-to-body, a single threshold is used between the nose of one animal and any body part of the other (except nose and tail base). Side-by-side and side-reverse-side are computed using two equal thresholds, measuring the distance between both noses and two tails in the former, and both nose-to-tail distances in the latter.

Of the individual traits, “look around” requires the animal to stand still (speed to be below a defined threshold) and the head to be moving (nose and ear speeds to be above a defined threshold). Finally, sniffing and wall climbing rely on the interaction of each animal with the arena (which can be detected automatically in certain settings, or indicated manually by the user using a GUI—graphical user interface—when creating a DeepOF project). An animal is annotated as sniffing the walls when speed is below a defined threshold, the distance between the nose and the wall is below a defined threshold, and the head is moving. Consequently, wall climbing is detected when the nose of an animal goes more than a certain threshold beyond the delimited arena. All mentioned thresholds can be specified (in millimeters) by the user. All analyses presented in this article were conducted with default values, which can be seen in Supplementary Table 1. Moreover, all annotations require a reported tracking likelihood of at least 0.85 on all involved body parts.

Annotation using pre-trained machine learning models. In the case of stopped and huddled, we trained a gradient boosting machine (scikit-learn, v1.2.0, default parameters) to detect the trait per frame, using a set of 26 variables including distances between body parts, speeds, and areas. Data were preprocessed by standardizing each animal’s trajectories independently (controlling for body size), and the training set as a whole. Furthermore, to deal with the imbalanced nature of the dataset (as only 8.48 % of the frames were positively labeled) we applied Synthetic Minority Over-sampling Technique (SMOTE)⁴⁹ to oversample the minority class (using imblearn v0.10.1⁵⁰).

Performance was then evaluated using a tenfold stratified cross-validation (to keep approximately the same number of positive labels in each validation fold) on a single dataset for model development and tested externally using a leave-one-dataset-out approach. Four independent datasets were used, collected in four different settings and across two different labs (see dataset details in Supplementary Table 2). Three of them (SI, OF, and SA) were tagged with manual labeling only, whereas the fourth (EX, obtained externally) combined manual labels and automatic pseudo-labeling using SimBA (Supplementary Fig. 2). The final classifier deployed with the latest version of DeepOF was then trained on a set of more than half a million labeled frames (567,367), coming from all four mentioned independent datasets, and global feature importance was obtained using SHAP (Shapley additive explanations).

After applying the annotators, a Kleinberg burst detection algorithm^{37,51} is applied to all predictions. This step smoothens the results by merging detections that are close in time (called bursts) and removing isolated detections, which an infinite hidden Markov model deems as noise. Moreover, rather than having a fixed detection window, the filter will be less likely to ignore isolated or less frequent events if they are far enough from higher frequency bursts but will be more prone to removing isolated events closer to a region where annotations are more frequent. In addition, it is important to notice that the annotators work independently, so more than one label can be assigned to an animal at a given time point (Fig. 1D).

Overall, while the provided behavioral set may not cover all scenarios, this out-of-the-box pipeline can be used to detect differences in behavior across experimental conditions without the need for further programming. More complex behaviors, involving user definition and labeling can thus be extracted using other available tools if required³⁷.

Graph representations

To analyze complex spatio-temporal data involving features such as coordinates, speed, and distances, the unsupervised pipeline within DeepOF can structure the variables as an annotated graph (Fig. 1E).

In this representation, each node is annotated with three values, corresponding to both coordinates of each body part, as well as their speeds. Edges are in turn annotated with distances between both connected body parts. The adjacency matrix describing connectivity is provided by DeepOF for top-down videos, but can also be defined by the user. Moreover, this representation can be extended to a multi-animal setting, where independent graph representations for each animal are connected through nose-to-nose, nose-to-tail, and tail-to-tail edges, allowing the models to incorporate relative distances between animals. It is worth mentioning that the provided representation works best when adjacent body parts are being tracked so that propagation through space is not too coarse. One of the main assumptions behind spatio-temporal graph embeddings is that connected body parts are sufficiently correlated in space, which may not be the case if too little tracking labels are included⁵².

Unsupervised deep embeddings with DeepOF

Unsupervised analysis of behavior was conducted using an integrated workflow within DeepOF, which enables both the deep embedding of animal trajectories and their clustering, to retrieve motion motifs that are consistent across time.

To this end, node and edge features (for either single or multiple animals) are processed using a sliding window across time, and standardized twice: once per animal, to remove size variability, and a second time on the entire training set.

The resulting data is then embedded using a deep clustering neural network architecture based on Variational Deep Embeddings^{53,54}, a deep clustering algorithm that can be adapted to sequential data. During training of the models, DeepOF minimizes the ELBO (evidence lower bound), represented in Eq. (1):

$$L_{\text{ELBO}}(x) = \mathbb{E}_{q(z,c|x)}[\log p(x|z)] - D_{\text{KL}}(q(z,c|x)||p(z,c)) \quad (1)$$

The first term corresponds to the reconstruction loss, which encourages the latent space (z) to represent the data (x) well over a set of clusters (c). The second term is the Kullback-Leibler divergence (D_{KL}) between a mixture-of-Gaussians prior ($p(z,c)$) and the variational posterior for each cluster ($q(z,c|x)$), which regularizes the embeddings to follow a mixture-of-Gaussians distribution where each component is associated with a particular behavior. A schematic overview of the model can be found in Fig. 1F.

Importantly, this loss function enforces a clustering structure directly in the latent space, removing the need for post-hoc clustering of the embeddings required by other available tools³³. This has several advantages, the main one being that the clustering structure back-propagates to the encoder during training, improving clustering performance⁵⁵.

The main contribution of the provided architecture lies however in the encoder-decoder layers, which are designed to handle spatio-temporal graph data (in which connectivity is static, but node and edge attributes change over time)⁵⁶. To accomplish this, features corresponding to each body part are first processed independently by a temporal block, which consists of a one-dimensional convolutional neural network (CNN) and two gated recurrent unit (GRU) layers. Subsequently, the outputs of these layers are passed by a spatial block, that shares information across adjacent body parts. This is accomplished using CensNet convolutions, a graph convolution architecture capable of embedding node and edge attributes at the same time⁵⁷. This allows DeepOF to take advantage of several data modalities related to motion with a single data structure as input.

Once the models are trained, cluster assignments are obtained as the argmax of the posterior distribution given the data, as described in Eq. (2):

$$q(c|x) = p(c|z) \equiv \frac{p(z)p(z|c)}{\sum_{c'=1}^K p(c')p(z|c')} \quad (2)$$

where $c' \in (1, K)$ is an iterator over all clusters in the model.

In practice, this unsupervised pipeline can retrieve consistent patterns of animal motion in a flexible, non-linear, and fully unsupervised way. Moreover, as body part speeds and distances can be naturally included, this workflow works even when critical body parts (such as the paws) are occluded, which makes it ideal for top-down videos.

In addition, DeepOF is capable of training multi-animal embeddings by using multi-animal graphs (see graph representations section above). When more than one animal is detected, DeepOF allows the user to control how much these embeddings should consider interactions between the animals over the multi-animal system. This is achieved with an L1 penalization over the node embeddings in the aforementioned CensNet layers: larger values will prime the models to prioritize animal interactions, whereas smaller values will increase the contribution of the individual behavior of each animal. All experiments included in this study used a moderated parameter (0.25) which allowed the model to consider both interactions and joint individual behaviors.

Unsupervised model training and hyperparameters

All unsupervised models used default values (as specified in DeepOF version 0.4.6). On each dataset, 10% of the available videos were used as a validation set to evaluate performance during training. Data were processed using sliding windows with a length matching the video frame rate of each dataset and stride of 1, mapping to eight-dimensional latent spaces. The training was conducted using the Nadam optimizer⁵⁸ (with a learning rate of 0.001 and gradient-based clipping of 0.75) over 150 epochs with early stopping based on the total validation loss and patience of 15 epochs. Upon training end, weights of the models are restored to those obtained in the best performing epoch using the same metric. The number of populated clusters over time, confidence in selected clusters (as the argmax of the produced soft counts), regularizers, and individual components of the loss function (see unsupervised deep embeddings with DeepOF section above) are tracked over time by DeepOF.

Global animal embeddings

Aside from embedding time points individually, global animal embeddings (where each data point corresponds to the trajectory of an entire animal rather than to a single time point) were obtained by constructing a k -dimensional vector with the time proportion each animal spent on each cluster, where k is the number of clusters in the given model.

Cluster number selection

For each dataset that was analyzed with the unsupervised pipeline, models ranging from 5 to 25 clusters were trained five times, resulting in a total of 120 models per explored setting. All model hyperparameters were set to DeepOF defaults (see section below and API documentation for additional details). Global animal embeddings were then used as input to a logistic regression classifier (scikit-learn, default parameters) aiming to discriminate CSDS from non-stressed animals. The model with the smallest number of clusters that reached a performance within one standard deviation of the global maximum across the whole range (in terms of the area under the ROC—receiver operating characteristic—curve) was selected for further processing.

Time binning and habituation quantification

A key aspect of DeepOF is that it allows for quantification of behavioral differences between cohorts over time in an unsupervised way. In this context, this is done by measuring the Wasserstein distance over time between the multivariate distributions describing global animal embeddings for CSDS and non-stressed animals.

By measuring this distance across a growing window, we can quantify how important additional information is to discriminate between conditions. This way, a peak in the distance curve would mark the point in time in which behavioral differences are maximized. In this study, we used a range between 10 and 600 s for each experiment, computing the Wasserstein distance between conditions every second. The time point at which the maximum was reached was selected as the optimal size for consecutive (non-overlapping) time bins. By reporting the behavioral distance along these bins, DeepOF can report behavioral habituation (which would involve behavioral differences between conditions decreasing over time).

Unsupervised cluster interpretation using Shapley additive explanations (SHAP)

When applying the unsupervised pipeline, and quantifying which features DeepOF deems relevant for the unsupervised models to determine the assignment of a given time segment to a given cluster, all obtained sequence-cluster mappings were analyzed using Shapley additive explanations^{59,60}.

To this end, a comprehensive set of 52 distinct features (111 for two-animal embeddings) was built to describe each sliding window in the training set, including mean values of distances, angles, speeds, and supervised annotators.

Gradient boosting machines (using Catboost v1.1.1⁶¹, which offers models specifically optimized for non-binary classification) were then trained to predict cluster labels from this set of statistics after normalization across the dataset and oversampling the minority class with the SMOTE algorithm⁴⁹. Performance is reported as the validation balanced accuracy across a 10-fold stratified cross-validation loop, and feature importance (global and for each cluster) is reported in terms of the average absolute SHAP values, obtained using a permutation explainer.

Animals for chronic social defeat stress experiments

Eight-week-old experimental male C57Bl/6N mice were bred in-house. The CD1 male mice (bred in-house) were used in the social avoidance and social interaction task as social conspecifics (CD1 animals were 4–6 weeks old) and as aggressors in the CSDS paradigm (CD1 animals were at least 16 weeks old). The study was conducted with male animals as a proof of principle, and for comparability to widely available data on chronic social defeat. All animals were housed in individually-ventilated cages (IVC; 30 cm × 16 cm × 16 cm connected by a central airflow system: Tecniplast, IVC Green Line—GM500) at least 2 weeks before the start of the experiment to allow acclimatization to the behavioral testing facility. All animals were kept under standard housing conditions; 12 h/12 h light-dark cycle (lights on at 7 a.m.), temperature 23 ± 1 °C, humidity 55%. Food (Altromin 1324, Altromin GmbH, Germany) and water were available *ad libitum*. All experimental procedures were approved by the committee for the Care and Use of Laboratory Animals of the government of Upper Bavaria, Germany. All experiments were in accordance with the European Communities Council Directive 2010/63/EU.

Chronic social defeat stress

At 2 months of age, male mice were randomly divided into the CSDS condition ($n = 30$) or the non-stressed condition (NS) ($n = 30$) (Supplementary Table 2, experiment code 1). The CSDS paradigm consisted of exposing the experimental C57Bl/6 N mouse to an aggressive CD1 mouse for 21 consecutive days, as previously described⁶². An additional

cohort (NS: $n = 30$, CSDS: $n = 33$, subdivided into susceptible animals $n = 9$, and resilient animals $n = 24$) was used to test the DeepOF social interaction classifiers on the resiliency and susceptibility division of the social avoidance ratio (Supplementary Table 2, experiment code 2). The prolonged 3-week CSDS paradigm was specifically chosen to elicit a more profound passive defeat phenotype, as originally reported by Kudryavtseva et al.¹³, and to allow multiple behavioral assessments under stress conditions. In short, the CD1 aggressor mice were trained and specifically selected on their aggression prior to the start of the experiment. The experimental mice were introduced daily to a novel CD1 resident's territory, who attacked and forced the experimental mouse into subordination. Defeat sessions lasted until the stress-exposed mouse received two bouts of attacks from the CD1 aggressor or at 5 min in the rare instances when two bouts were not achieved within this duration. Animal health was monitored throughout the experiment to ensure that any minor injuries healed prior to the subsequent defeat session. Between daily defeats, stressed mice were housed in the resident's home cage but physically separated from the resident by a see-through, perforated mesh barrier, allowing sensory exposure to the CD1 aggressor mouse while preventing further attacks. The defeat time of day was randomized between 11 a.m. and 6 p.m. to avoid habituation and anticipatory behaviors in defeated mice. NS mice were single-housed in the same room as the stressed mice. All animals were handled daily and weighed every 3–4 days. Behavioral testing was performed after 14 days of the defeat paradigm, where behavior was observed in the morning and the defeat continued in the afternoon. The animals were sacrificed a day after the CSDS ended under deep isoflurane anesthesia by decapitation, which was at 3 months of age. Then, the adrenals were obtained, and the relative adrenal weight was calculated by dividing the adrenal weight by the body weight before sacrifice.

Behavioral testing

Behavioral tests were performed between 8 a.m. and 11 a.m. in the same room as the housing facility. On day 15 of the CSDS paradigm, the animals were tested on the social avoidance (SA) task, while on day 16, the animals were tested on the combined open field (OF) and social interaction (SI) task. The SA task was analyzed using the automated video-tracking software AnyMaze 6.33 (Stoelting, Dublin, Ireland), whereas the OF and SI tasks were analyzed using DeepLabCut 2.2b7 for pose estimation^{29,30}, after which DeepOF module version 0.4.6 was used for preprocessing, supervised, and unsupervised analyses of behavior.

Social avoidance

The SA task was performed in a square OF arena (50×50 cm) to observe the social behavioral profile after CSDS, as well-established in previous studies^{13,62–64}. The SA task consisted of two phases: the non-social stimulus phase and the social stimulus phase. During the non-social stimulus phase, which was the first 2.5 min, the experimental mouse was allowed to freely explore the OF arena with a small empty wired mesh cage against the wall of the OF. Then, the empty wired mesh cage was replaced with a wired mesh cage including a trapped unfamiliar young CD1 mouse (4–6 weeks old). During the following 2.5 min, the social-stimulus phase, the experimental mouse could freely explore the arena again. The SA-ratio was calculated by calculating the amount of time spent with the social stimulus, which was then divided by the time spent with the non-social stimulus. The identification of CSDS susceptibility and resiliency was obtained using a SA-ratio score of lower than “1” for susceptible animals, and an SI-ratio score higher than “1” for resilient animals.

Open field and social interaction task

The OF and SI tasks were performed in a round OF arena (diameter of 38 cm). The bottom of the arena was covered in sawdust material to

minimize the cross-over effects of stress and anxiety by the novel environment. First, the OF task was performed, during which the experimental animal was allowed to freely explore the arena for 10 min. Subsequently, for the SI task, an unfamiliar young CD1 (4–6 weeks old) was introduced inside the arena and both animals were allowed to freely explore the arena for 10 min. The DeepOF module can identify five behavioral traits during the single animal OF task, which include wall-climbing, stopped-and-huddled, look-around, sniffing, and speed (locomotion), whereas in the SI task, all behavioral traits can be identified (Fig. 1C). During the analysis, the 10 min OF and SI tasks were analyzed in the total duration of the behavioral classifiers, and in time bins of 2.5 min to match the time frame in the SA task.

Z-score stress physiology and social interaction calculation

The Z-scores combine the outcome of multiple tests via mean normalization and provide an overall score for the related behavior of interest. Z-scores were calculated as described previously⁶⁵. The Z-score indicates for every observation (X), the number of standard deviations (σ) above or below the mean of the control group (μ). This means that for each individual observation Eq. (3) is calculated:

$$Z = \frac{X - \mu}{\sigma} \quad (3)$$

Then, the obtained values need to be corrected for the directionality, such that an increased score will reflect the increase of the related behavior of interest. This means that per test, the scores were either already correct or were adjusted in the correct directionality by multiplying with “–1”. Finally, to calculate the final z-score, the different z-scores per test were combined and divided by the total number of tests, as in Eq. (4).

$$Z_{total} = \frac{\sum_i z_{test_i}}{\text{Number of tests}} \quad (4)$$

The Z-score analysis of stress physiology is based on the relative adrenal weight and the body weight at day 21 of the CSDS, which are both strongly influenced by CSDS exposure¹². The directionality of both tests did not require additional adjustment. Then, the Z-score of SI was calculated based on five DeepOF behavioral classifiers from the C57Bl/6N mouse, which were B-look-around, B-speed, B-huddle, B-nose-to-tail, and B-nose-to-body. The directionality was adjusted for B-speed, B-nose-to-tail, and B-nose-to-body.

Behavioral entropy calculation

Shannon's entropy of the behavioral cluster space was obtained directly using DeepOF, as a measure of how predictable the sequence of behaviors expressed by a given animal is. To accomplish this, DeepOF obtains transition matrices across clusters using the unsupervised cluster assignments per animal. Stationary distributions for each transition matrix are then obtained by simulation through the matrices until convergence, and Shannon's entropy is computed for each stationary distribution. Entropy scores obtained for NS and CSDS animals were then compared. Overall entropy scores were also compared to the stress physiology Z-score for validation purposes.

External dataset for validation of the DeepOF huddle classifier

An additional experiment was performed using different conditions and behavioral set-up, to assess the transferability of the DeepOF huddle classifier (Supplementary Table 2, experiment code 3) to data produced by a different lab. 12 weeks old C57BL/6J mice ($n = 24$, purchased from the Jackson Laboratory (catalog number 000664), Bar Harbor, ME, USA) were paired in a home-cage environment (19×19 cm) with 12 weeks old ovariectomized CFW female mice

(purchased from Charles River Laboratories (catalog number 024), Wilmington, MA, USA) and were allowed to freely explore each other for 1.5 min. The animals were housed under standard laboratory conditions with a 12 h light–dark cycle (lights on from 07:00 to 19:00), temperature $22 \pm 1^\circ\text{C}$, humidity 50%, in clear Plexiglas cages ($19 \times 29 \times 13$ cm) with unrestricted access to food (Purina Laboratory Rodent Diet 5001) and water. Procedures were approved by the McLean Hospital Institutional Animal Care and Use Committee and complied with the National Institutes of Health guidelines.

Statistics

Statistical analyses and graphs were made in RStudio (R 4.1.1⁶⁶) and python (v 3.9.13). All data were used for further statistical analysis unless stated otherwise. During the DeepLabCut tracking, seven animals were excluded due to technical difficulties (four NS and three CSDS were excluded). Statistical assumptions were then checked, in which the data were tested for normality using the Shapiro-Wilk test and QQ-plots and for heteroscedasticity using Levene's test. Data that violated these assumptions were analyzed using non-parametric tests. The time-course data was analyzed using the two-way ANOVA (parametric) or Kruskal-Wallis test (non-parametric) with time (days) as a within-subject factor and condition (NS vs. CSDS) as a between-subject factor, further posthoc analysis was performed using the Benjamini-Hochberg (BH) test (parametric) or the Wilcoxon test (non-parametric). P-values were adjusted for multiple testing using the Benjamini-Hochberg (BH) method. Three-group comparisons were analyzed using the one-way ANOVA (parametric) or Kruskal-Wallis test (non-parametric), and further posthoc analysis was performed using the BH test (parametric) or the Wilcoxon test (non-parametric). Two-group comparisons were analyzed using independent samples *t*-tests (parametric), Welch's tests (data are normalized but heteroscedastic), or Wilcoxon tests (non-parametric). Correlation analyses were performed using the Pearson correlation coefficient; outliers deviating more than 5 standard deviations from a fitted linear model were excluded from the analysis. The timeline and bar graphs are presented as mean \pm standard error of the mean. Data was considered significant at $p < 0.05$ (*), with $p < 0.01$ (**), $p < 0.001$ (***), $p < 0.0001$ (****).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The authors declare that data supporting the findings of this study are available within the Article and Supplementary Information. Source data are provided with this paper.

Code availability

All data and the accompanying code to perform the analyses and creating the figures are available for download via the Max Planck DataShare services. The most recent version of DeepOF is hosted in a GitHub repository, and a Zenodo release of the version used in this manuscript (v0.4.6) is found under <https://doi.org/10.5281/zenodo.8013401>. The most recent stable version of DeepOF is available in PyPI. Full documentation and tutorials are available on read the docs.

References

- World Health Organisation. *Depression and Other Common Mental Disorders: Global Health Estimates*. Geneva: World Health Organization (2017).
- Cipriani, A. et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *Focus* **16**, 420–429 (2018).
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. <https://doi.org/10.1176/appi.books.9780890425596> (American Psychiatric Association, 2013).
- Addabbo, T., Sarti, E. & Sciulli, D. Healthy life, social interaction and disability. *Qual. Quant.* **50**, 2609–2623 (2016).
- Giordano, G. N. & Lindstrom, M. The impact of changes in different aspects of social capital and material conditions on self-rated health over time: a longitudinal cohort study. *Soc. Sci. Med.* **70**, 700–710 (2010).
- Manchia, M. et al. The impact of the prolonged COVID-19 pandemic on stress resilience and mental health: a critical review across waves. *Eur. Neuropsychopharmacol.* **55**, 22–83 <https://doi.org/10.1016/j.euroneuro.2021.10.864> (2021).
- Santomauro, D. F. et al. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. *Lancet* **398**, 1700–1712 [https://doi.org/10.1016/S0140-6736\(21\)02143-7](https://doi.org/10.1016/S0140-6736(21)02143-7) (2021).
- Gururajan, A., Reif, A., Cryan, J. F. & Slattery, D. A. The future of rodent models in depression research. *Nat. Rev. Neurosci.* **20**, 686–701 (2019).
- von Mücke-Heim, I.-A. et al. Introducing a depression-like syndrome for translational neuropsychiatry: a plea for taxonomical validity and improved comparability between humans and mice. *Mol. Psychiatry* **28**, 329–340 <https://doi.org/10.1038/s41380-022-01762-w> (2022).
- Golden, S. A., Covington, H. E., Berton, O. & Russo, S. J. A standardized protocol for repeated social defeat stress in mice. *Nat. Protoc.* **6**, 1183–1191 (2011).
- Russo, S. J. & Nestler, E. J. The brain reward circuitry in mood disorders. *Nat. Rev. Neurosci.* **14**, 609–625 (2013).
- Hollis, F. & Kabbaj, M. Social defeat as an animal model for depression. *ILAR J.* **55**, 221–232 (2014).
- Kudryavtseva, N. N., Bakshantovskaya, I. V. & Koryakina, L. A. Social model of depression in mice of C57BL/6J strain. *Pharm. Biochem. Behav.* **38**, 315–320 (1991).
- Donahue, R. J., Muschamp, J. W., Russo, S. J., Nestler, E. J. & Carlezon, W. A. Effects of striatal δ fosb overexpression and ketamine on social defeat stress-induced anhedonia in mice. *Biol. Psychiatry* **76**, 550–558 (2014).
- Iñiguez, S. D. et al. Social defeat stress induces a depression-like phenotype in adolescent male c57BL/6 mice. *Stress* **17**, 247–255 (2014).
- Yoshida, K. et al. Chronic social defeat stress impairs goal-directed behavior through dysregulation of ventral hippocampal activity in male mice. *Neuropsychopharmacology* **46**, 1606–1616 (2021).
- Krishnan, V. et al. Molecular adaptations underlying susceptibility and resistance to social defeat in brain reward regions. *Cell* **131**, 391–404 (2007).
- Toth, I. & Neumann, I. D. Animal models of social avoidance and social fear. *Cell Tissue Res.* **354**, 107–118 (2013).
- Sturman, O. et al. Deep learning-based behavioral analysis reaches human accuracy and is capable of outperforming commercial solutions. *Neuropsychopharmacology* **45**, 1942–1952 (2020).
- Hånell, A. & Marklund, N. Structured evaluation of rodent behavioral tests used in drug discovery research. *Front. Behav. Neurosci.* **8**, 252 (2014).
- Goodwin, N. L., Nilsson, S. R. O. & Golden, S. A. Rage against the machine: advancing the study of aggression ethology via machine learning. *Psychopharmacology* **237**, 2569–2588 (2020).
- Pearson, B. L., Defensor, E. B., Blanchard, D. C. & Blanchard, R. J. C57BL/6J mice fail to exhibit preference for social novelty in the three-chamber apparatus. *Behav. Brain Res.* **213**, 189–194 (2010).
- Lorsch, Z. S. et al. Computational analysis of multidimensional behavioral alterations after chronic social defeat stress. *Biol. Psychiatry* **89**, 920–928 (2021).

24. Marks, M., et al Deep-learning based identification, pose estimation and end-to-end behavior classification for interacting primates and mice in complex environments. (2021).
25. Pereira, T. D. et al. Fast animal pose estimation using deep neural networks. *Nat. Methods* **16**, 117–125 (2019).
26. Hsu, A. I. & Yttri, E. A. B-SOiD, an open-source unsupervised algorithm for identification and fast prediction of behaviors. *Nat. Commun.* **12**, 5188 (2021).
27. Segalin, C. et al. The mouse action recognition system (MARS) software pipeline for automated analysis of social behaviors in mice. *Elife* **10**, e63720 (2021).
28. de Chaumont, F. et al. Real-time analysis of the behaviour of groups of mice via a depth-sensing camera and machine learning. *Nat. Biomed. Eng.* **3**, 930–942 (2019).
29. Mathis, A. et al. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **21**, 1281–1289 (2018).
30. Lauer, J. et al. Multi-animal pose estimation, identification and tracking with DeepLabCut. *Nat. Methods* **19**, 496–504 (2022).
31. Wiltschko, A. B. et al. Mapping sub-second structure in mouse behavior. *Neuron* **88**, 1121–1135 (2015).
32. Kabra, M., Robie, A. A., Rivera-Alba, M., Branson, S. & Branson, K. JAABA: interactive machine learning for automatic annotation of animal behavior. *Nat. Methods* **10**, 64–67 (2013).
33. Luxem, K. et al. Identifying behavioral structure from deep variational embeddings of animal motion. *Commun. Biol.* **5**, 1267 (2022).
34. Schmidt, M. V. & Koutsouleris, N. Promises and pitfalls of the new era of computational behavioral neuroscience. *Biol. Psychiatry* **89**, 845–846 (2021).
35. Miranda, L., Bordes, J., Gasperoni, S. & Lopez, J. P. Increasing resolution in stress neurobiology: from single cells to complex group behaviors. *Stress* **26**, 2186141 (2023).
36. Miranda, L., Bordes, J., Pütz, B., Schmidt, M. V. & Müller-Myhsok, B. DeepOF: a Python package for supervised and unsupervised pattern recognition in mice motion tracking data. *J. Open Source Softw.* **8**, 5394 (2023).
37. Nilsson, S. R. O. et al. Simple Behavioral Analysis (SimBA)—an open source toolkit for computer classification of complex social behaviors in experimental animals. *bioRxiv* 2020.04.19.049452 <https://doi.org/10.1101/2020.04.19.049452> (2020).
38. Lopez, J. P. et al. Single-cell molecular profiling of all three components of the HPA axis reveals adrenal ABCB1 as a regulator of stress adaptation. *Sci. Adv.* **7**, eabe4497 (2021).
39. Gheusi, G., Bluthé, R.-M., Goodall, G. & Dantzer, R. Social and individual recognition in rodents: methodological aspects and neurobiological bases. *Behav. Process.* **33**, 59–87 (1994).
40. Brudzynski, S. M. Ethotransmission: communication of emotional states through ultrasonic vocalization in rats. *Curr. Opin. Neurobiol.* **23**, 310–317 (2013).
41. Hamasato, E. K., Lovelock, D., Palermo-Neto, J. & Deak, T. Assessment of social behavior directed toward sick partners and its relation to central cytokine expression in rats. *Physiol. Behav.* **182**, 128–136 (2017).
42. Rogers-Carter, M. M., Djerdjaj, A., Culp, A. R., Elbaz, J. A. & Christianson, J. P. Familiarity modulates social approach toward stressed conspecifics in female rats. *PLoS One* **13**, e0200971 (2018).
43. Morel, C. et al. Midbrain projection to the basolateral amygdala encodes anxiety-like but not depression-like behaviors. *Nat. Commun.* **13**, 1532 (2022).
44. Ye, S., Mathis, A. & Mathis, M. W. Panoptic animal pose estimators are zero-shot performers. (2022).
45. Schneider, S., Lee, J. H. & Mathis, M. W. Learnable latent embeddings for joint behavioral and neural analysis. *Nature* **617**, 360–368 (2022).
46. Sun, J. J. et al. The multi-agent behavior dataset: mouse dyadic social interactions. (2021).
47. Savitzky, A. & Golay, M. J. E. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **36**, 1627–1639 (1964).
48. Buitinck, L. et al. API design for machine learning software: experiences from the scikit-learn project. (2013).
49. Chawla, N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P. SMOTE: synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **16**, 321–357 (2002).
50. Lemaitre, G., Nogueira, F. & Aridas, C. K. Imbalanced-learn: a Python toolbox to tackle the curse of imbalanced datasets in machine learning. *jmlr. org.* **18**, 1–5 (2017).
51. Kleinberg, J. Bursty and hierarchical structure in streams. In: *Proc. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 91–101. <https://doi.org/10.1145/775060.775061> (2002).
52. Zhou, J. et al. Graph neural networks: a review of methods and applications. *AI Open* **1**, 57–81 (2020).
53. Jiang, Z., Zheng, Y., Tan, H., Tang, B. & Zhou, H. Variational deep embedding: an unsupervised and generative approach to clustering. In: *Proc. Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17)* (2016).
54. Manduchi, L. et al. A Deep variational approach to clustering survival data. (2021).
55. Lafabregue, B., Weber, J., Gançarski, P. & Forestier, G. End-to-end deep representation learning for time series clustering: a comparative study. *Data Min. Knowl. Discov.* **36**, 29–81 (2022).
56. Yu, B., Yin, H. & Zhu, Z. Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting. In: *Proc. Twenty-Seventh International Joint Conference on Artificial Intelligence*. 3634–3640. <https://doi.org/10.24963/ijcai.2018/505> (2018).
57. Jiang, X., Zhu, R., Ji, P. & Li, S. Co-Embedding of Nodes and Edges With Graph Neural Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **45**, 7075–7086 (2023).
58. Dozat, T. Incorporating Nesterov Momentum into Adam. Preprint at (2016).
59. Goodwin, N. L., Nilsson, S. R. O., Choong, J. J. & Golden, S. A. Toward the explainability, transparency, and universality of machine learning for behavioral classification in neuroscience. *Curr. Opin. Neurobiol.* **73**, 102544 (2022).
60. Lundberg, S. & Lee, S.-I. A Unified Approach to Interpreting Model Predictions. (2017).
61. Dorogush, A. V., Ershov, V. & Gulin, A. CatBoost: gradient boosting with categorical features support. (2018).
62. Wagner, K. V. et al. Differences in FKBP51 regulation following chronic social defeat stress correlate with individual stress sensitivity: influence of paroxetine treatment. *Neuropsychopharmacology* **37**, 2797–2808 (2012).
63. Berton, O. et al. Essential role of BDNF in the mesolimbic dopamine pathway in social defeat stress. *Science (1979)* **311**, 864–868 (2006).
64. Tsankova, N. M. et al. Sustained hippocampal chromatin regulation in a mouse model of depression and antidepressant action. *Nat. Neurosci.* **9**, 519–525 (2006).
65. Guilloux, J.-P., Seney, M., Edgar, N. & Sibille, E. Integrated behavioral z-scoring increases the sensitivity and reliability of behavioral phenotyping in mice: Relevance to emotionality and sex. *J. Neurosci. Methods* **197**, 21–31 (2011).
66. R: The R Project for Statistical Computing. <https://www.r-project.org/>.
67. Zhao, Y. et al. Social rank-dependent effects of testosterone on huddling strategies in mice. *iScience* **26**, 106516 (2023).

Acknowledgements

The authors thank the DeepLabCut development team for creating and maintaining the DeepLabCut software. In addition, the authors thank Margherita Springer for the language editing of the manuscript and Max Pöhlmann for the design of the mouse illustrations in Fig. 2a. This study is supported by the Kids2Health grant of the Federal Ministry of Education and Research (01GL1743C, M.V.S.), and the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement (813533; L.M.).

Author contributions

JB and MVS conceived the study. L.M. wrote the DeepOF module, with primary technical assessment from F.A., B.P., and B.M.M. J.B. and M.R. performed the experiments. L.M.B., L.v.D., C.E., L.D., and S.M. assisted with the experiments. J.B. and L.M. analyzed the data and wrote the first version of the manuscript. B.P. worked on figure design. J.B. created the mouse illustrations in Fig. 1. S.N., J.H., E.L.N., and K.J.R. assisted with manual behavioral data tracking and analysis for data benchmarking purposes. All authors contributed to the revision of the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-40040-3>.

Correspondence and requests for materials should be addressed to Bertram Müller-Myhsok or Mathias V. Schmidt.

Peer review information *Nature Communications* thanks David Slattery, Johannes Bohacek and the other, anonymous, reviewer for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

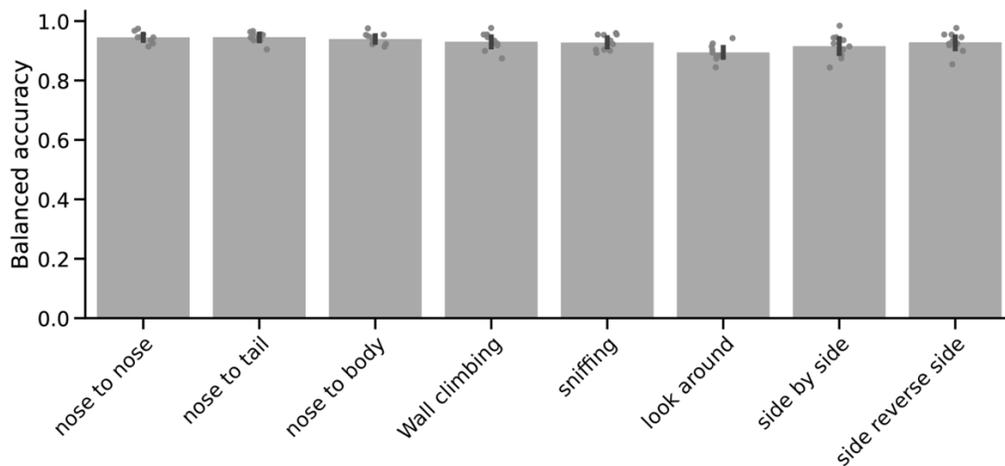
Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

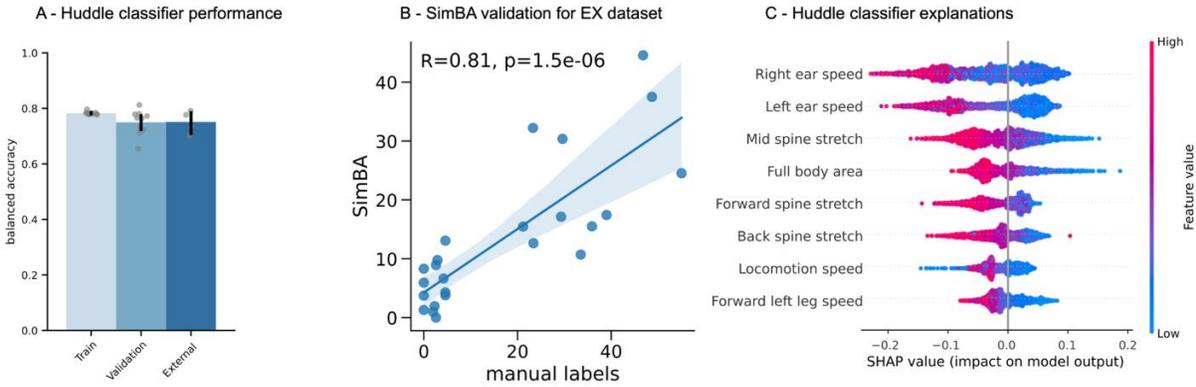
© The Author(s) 2023

Automatically annotated motion tracking identifies a distinct social behavioral profile following chronic social defeat stress

Supplemental material



Supplemental figure 1. Validation of rule-based annotated behaviors. 10 out of 53 videos were manually labeled for all annotators (excluding stopped-and-huddled, see supplemental figure 2) using the Colabeler software (v2.0.4). Balanced accuracy between manual labels and predicted binary outcomes (presence or absence of a given trait at a given time) is reported. Bars represent the mean \pm standard deviation across all 10 videos (N=10). Source data are provided as a Source Data file.



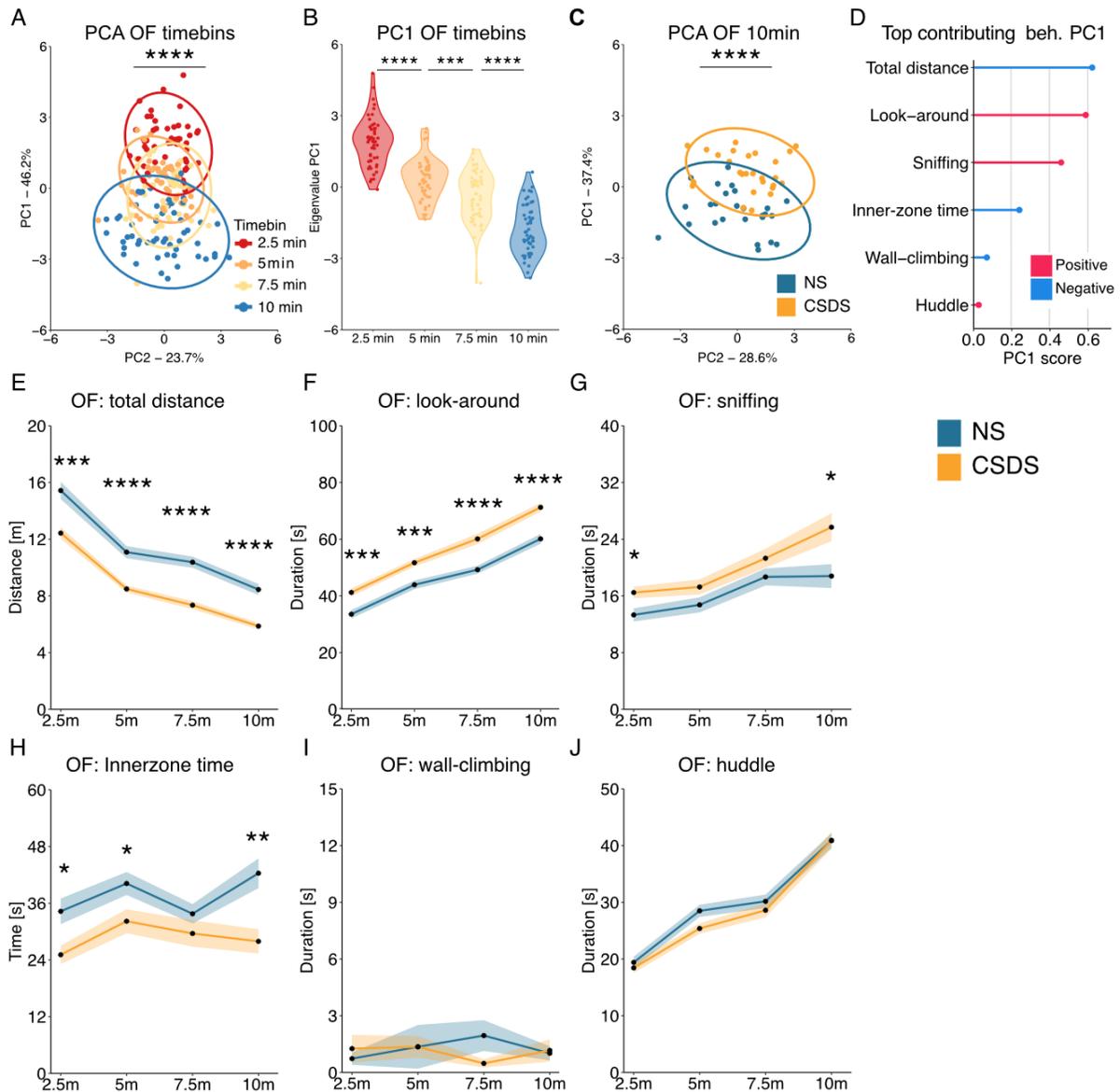
Supplemental figure 2. Validation of stopped-and-huddled classifier. A) Bar charts (mean \pm standard deviation) showing balanced accuracy performance for the huddle classifier provided with the supervised pipeline within DeepOF. A total of 567367 video frames were either manually labeled (for the SI, OF, and SA datasets) or pseudo-labeled using SimBA (EX dataset) for the stopped-and-huddled trait using the labeling tool provided with SimBA v1.31.1. Labelling was conducted in four independent datasets (SI, OF, SA, and EX; see the animals' section in materials and methods for details), and two validation tasks were conducted, marked as "Validation" and "External" respectively. First, a 10-fold stratified cross-validation loop was executed within the SI dataset (which has the most labels, see supplemental table 2 for details), to test for overfitting and generalization within a single dataset. Balanced accuracy results were 0.78 ± 0.005 and 0.75 ± 0.046 for the training and validation sets, respectively (N=10). Second, a leave-one-dataset out cross-validation was conducted across all four datasets, to test whether the model can generalize to novel settings (different bedding, different arenas, different labs). A balanced accuracy of 0.75 ± 0.04 was reported (N=4). B) SimBA validation of the classifiers used for pseudo-labelling in the external dataset. Correlation between total behavior duration (in seconds) in manual and predicted labels shown for all 24 videos (N=24). Both sets show a Pearson correlation coefficient $\rho=0.81$, which significantly deviates from zero (p -value= $1.5e-6$). Error bands represent the 95% confidence interval. C) SHAP analysis of the deployed model (trained in the whole dataset, with all concatenated four sites). The top 8 features are displayed of a total of 26 features including distances between body parts, speeds, and areas. Results show low head movement, low spine stretch, low body area, and low locomotion speed as the most important features for the model, which goes in line with the accepted definition of the behavior. Source data are provided as a Source Data file.

Supplemental table 1. Default thresholds used by the annotation pipeline in DeepOF

Annotated trait	Rule	Default threshold in DeepOF
Nose-to-nose	Nose to nose distance	< 25 mm
Nose-to-tail	Nose to tail distance	< 25 mm
Nose-to-body	Nose to any other body part	< 25 mm
Side-by-side	Nose to nose distance	< 45 mm
	Tail to tail distance	< 45 mm
Side-reverse-side	Nose to tail distance	< 45 mm
Wall-climbing	Nose reach beyond walls	> 10 mm
Sniffing	Nose distance to object	< 10 mm
	Nose speed	> 50 mm/s
	Locomotion speed	< 50 mm/s
Look-around	Locomotion speed	< 50 mm/s
	Nose speed	> 50 mm/s

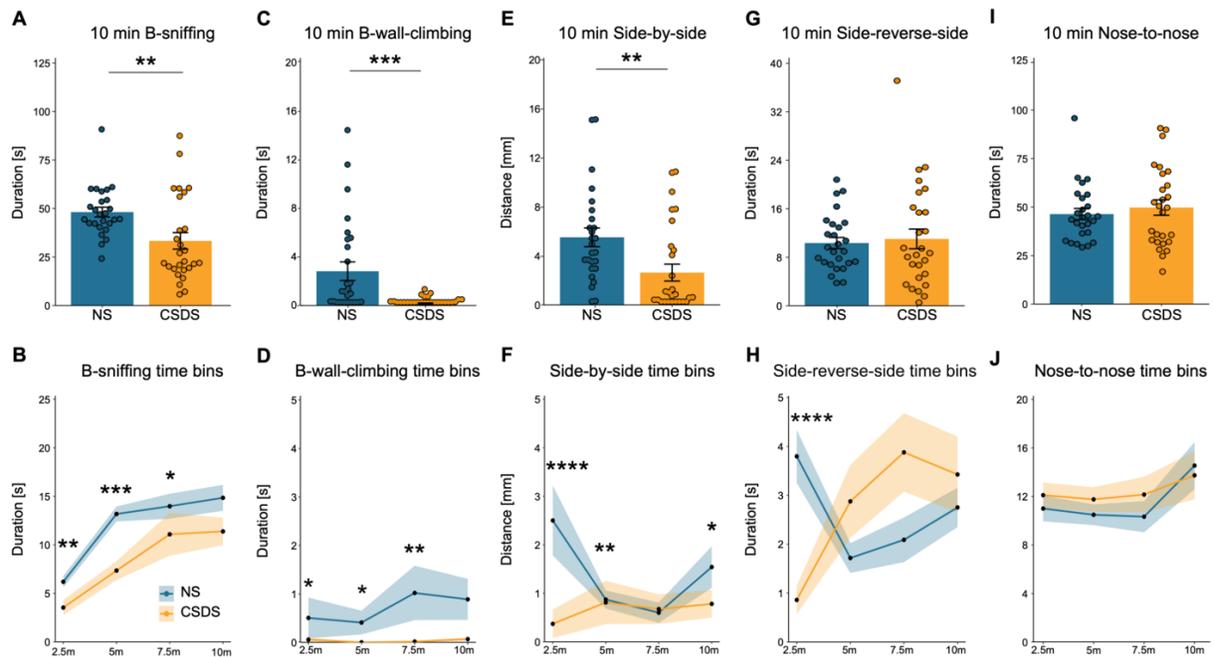
Supplemental table 2. Datasets used in the current study

Dataset name	Experiment code	Number of videos	Frame rate	Video length	Labeled frames (stopped-and-huddled)	Prevalence (stopped-and-huddled)
Social interaction (SI)	1	53	25	10 min 15000 frames	299.350	10.83%
Open field (OF)	1	53	25	10 min 15000 frames	179.979	2.75%
Social avoidance (SA)	1	120	13	2.5 min 1950 frames	22.488	4.36%
Social interaction for SA resiliency (figure S6)	2	64	30	10 min 18000 frames	0	-
Social interaction (external)	3	20	30	1.5 min 2730 frames	65.550	14.95%
Total	-	310	-	-	567.367	8.49%

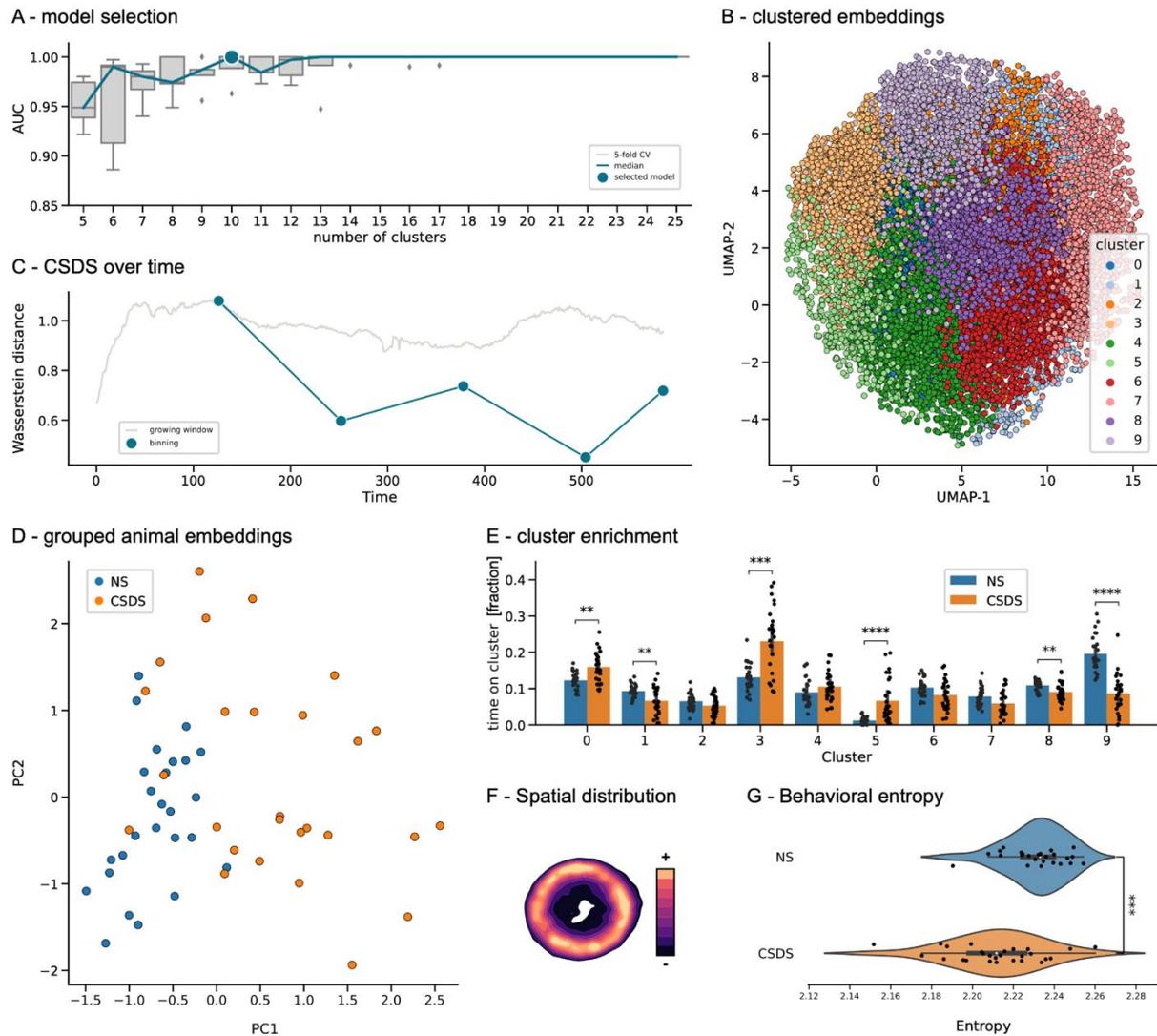


Supplemental figure 3. DeepOF behavioral classifiers in the open field task. A) The OF PCA time bins show a significant main effect (one-way ANOVA: $F(3,208)=129.12$, $p=2.97e-47$). B) Benjamini-hochberg (BH) posthoc shows that the time bins are significantly different from each other (2.5vs5, $p=3.93e-14$; 5vs7.5, $p=0.0003$, 7.5vs10, $p=3.1e-12$). C) The 10min OF PCA analysis shows a significant difference between conditions; independent samples t -test: $T(51)=-7.23$, $p=2.37e-9$. Data consisted of all the individual DeepOF behavioral classifiers, as listed in Figure 1C. D) The ranked behaviors on the PC1 using the corresponding rotated loading scores. E) The total distance was lower in CSDS animals; posthoc BH: 2.5 min $T(51)=16.89$, $p=0.0001$, 5 min $T(51)=28.28$, $p=3.13e-6$, 7.5 min $T(51)=39.59$, $p=2.86e-7$, and 10 min $T(51)=33.77$, $p=8.1e-7$. Two-way ANOVA on condition: $F(1,208)=92.586$, $p=2.31e-18$, time: $F(1,208)=265.77$, $p=4.85e-39$, condition \times time: $F(1,208)=0.10$, $p=0.75$). F) Look-around was higher in CSDS

animals; posthoc BH: 2.5 min ($T(51)=14.08$, $p=0.0004$), 5 min ($T(51)=14.84$, $p=0.0004$), 7.5 min ($T(51)=21.65$, $p=4.7e-5$), and 10 min ($T(51)=23.25$, $p=4.7e-5$). Two-way ANOVA on condition: $F(1,208)=74.04$, $p=1.9e-15$, time: $F(1,208)=356.65$, $p=5.4e-47$, condition×time: $F(1,208)=1.90$, $p=0.17$). G) Sniffing was higher in CSDS animals for the 2.5- and 10 min time bins; posthoc Wilcoxon: $W=199.5$, $p=0.023$; $W=210$, $p=0.023$, respectively. The 5- and 7.5 min were not altered ($W=258$, $p=0.13$, and $W=307$, $p=0.44$, respectively). Kruskal-Wallis test 2.5 min: $H(1)=7.27$, $p=0.024$, 5 min: $H(1)=2.74$, $p=0.13$, 7.5 min: $H(1)=0.6$, $p=0.43$, and 10 min: $H(1)=6.29$, $p=0.024$. H) The inner zone time was lowered in CSDS animals for the 2.5, 5, and 10 min time bins; posthoc BH: $T(51)=7.70$, $p=0.016$, $T(51)=5.16$, $p=0.036$, $T(51)=12.74$, $p=0.0032$, respectively). The 7.5 min was not altered ($p=0.24$). Two-way ANOVA on condition: $F(1,208)=24.04$, $p=1.9e-6$, time: $F(1,208)=2.07$, $p=0.15$, condition×time: $F(1,208)=0.53$, $p=0.47$. I) Climbing did not reveal any difference using the Kruskal-Wallis test. J) Huddle did not reveal any difference using the Kruskal-Wallis test. The PCA graphs are provided with a 95% confidence ellipse and all individual samples as points. Further PC1 analyses are represented with a violin plot and all individual samples as points. The timeline graphs are presented as mean \pm standard error of the mean. $N=26$ for NS and $n=27$ for CSDS in panels A-J. Source data are provided as a Source Data file.

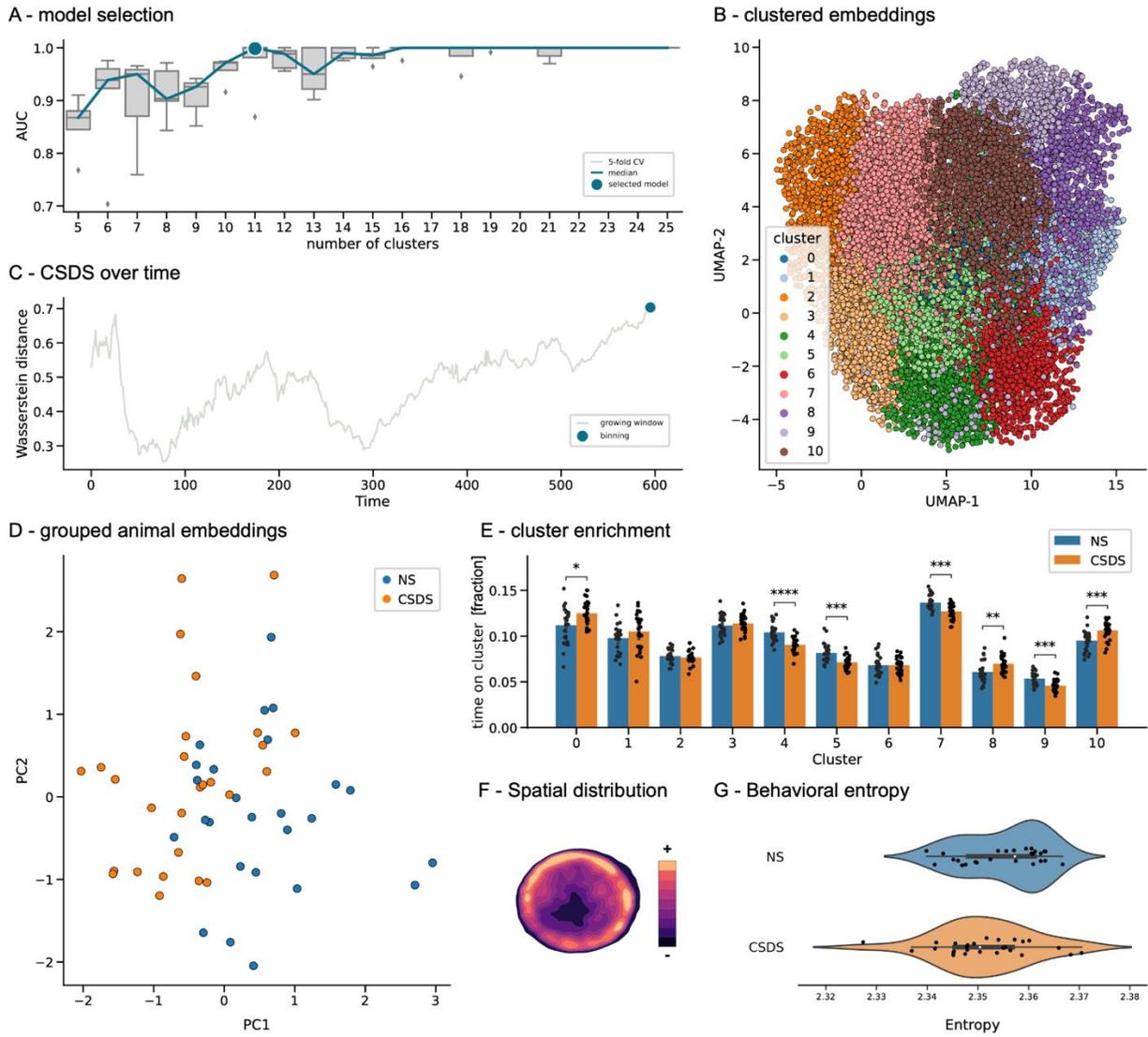


Supplemental Figure 4. DeepOF other behavioral classifiers in the social interaction task for 10 min duration. A) B-sniffing is lower in CSDS animals. Independent samples *t*-test: $T(51)=2.99$, $p=0.004$. B) Wilcoxon posthoc analysis revealed that B-sniffing was lower in CSDS animals for the 2.5 min ($W=538$, $p=0.002$), 5 min ($W=576$, $p=0.0003$), and 7.5 min ($W=499$, $p=0.012$), but not the 10 min ($W=456$, $p=0.06$). Kruskal-Wallis test: 2.5 min: $p=0.002$, 5 min: $p=0.0003$, 7.5 min: $p=0.012$, and 10 min: $p=0.06$. C) B-wall-climbing is lower in stressed animals. Wilcoxon test: $W=540$, $p=0.0004$. D) Wilcoxon posthoc analysis revealed that B-wall-climbing was lower in stressed animals for the 2.5 min ($W=441$, $p=0.03$), the 5 min ($W=435$, $p=0.03$), and the 7.5 min ($W=506$, $p=0.002$), but not the 10 min ($W=393$, $p=0.37$). Kruskal-Wallis test: 2.5 min: $p=0.03$, 5 min: $p=0.03$, 7.5 min: $p=0.002$, and 10 min: $p=0.37$. E) Side-by-side is lower in CSDS animals. Wilcoxon test: $W=522.5$, $p=0.0023$. F) Wilcoxon posthoc analysis revealed that Side-by-side was lower in CSDS animals for the 2.5 min ($W=581$, $p=5.48e-5$), the 5 min ($W=521.5$, $p=0.003$), and the 10 min ($W=491.5$, $p=0.02$), but not the 7.5 min ($W=405$, $p=0.32$). Kruskal-Wallis test: 2.5 min: $p=5.28e-5$, 5 min: $p=0.003$, 7.5 min: $p=0.32$, and 10 min: $p=0.02$. G) Side-reverse-side is not altered between conditions. Wilcoxon test: $W=365$, $p=0.81$. H) Wilcoxon posthoc analysis revealed that Side-reverse-side was lower in CSDS animals for the 2.5 min time bin ($W=628$, $p=3.36e-6$), but not the 5-, 7.5-, and 10 min time bins ($W=337.5$, $p=1$; $W=292.5$, $p=0.60$; $W=351$, $p=1$, respectively). Kruskal-Wallis test: 2.5 min: $p=3.21e-6$, 5 min: $p=1$, 7.5 min: $p=0.60$, and 10 min: $p=1$. I) Nose-to-nose is not altered between conditions. Wilcoxon test: $W=326$, $p=0.67$. J) No further significant differences were observed in the Nose-to-nose time bins. The timeline and bar graphs are presented as mean \pm standard error of the mean and all individual samples as points. $N=26$ for NS and $N=27$ for CSDS in panels A-J. Source data are provided as a Source Data file.



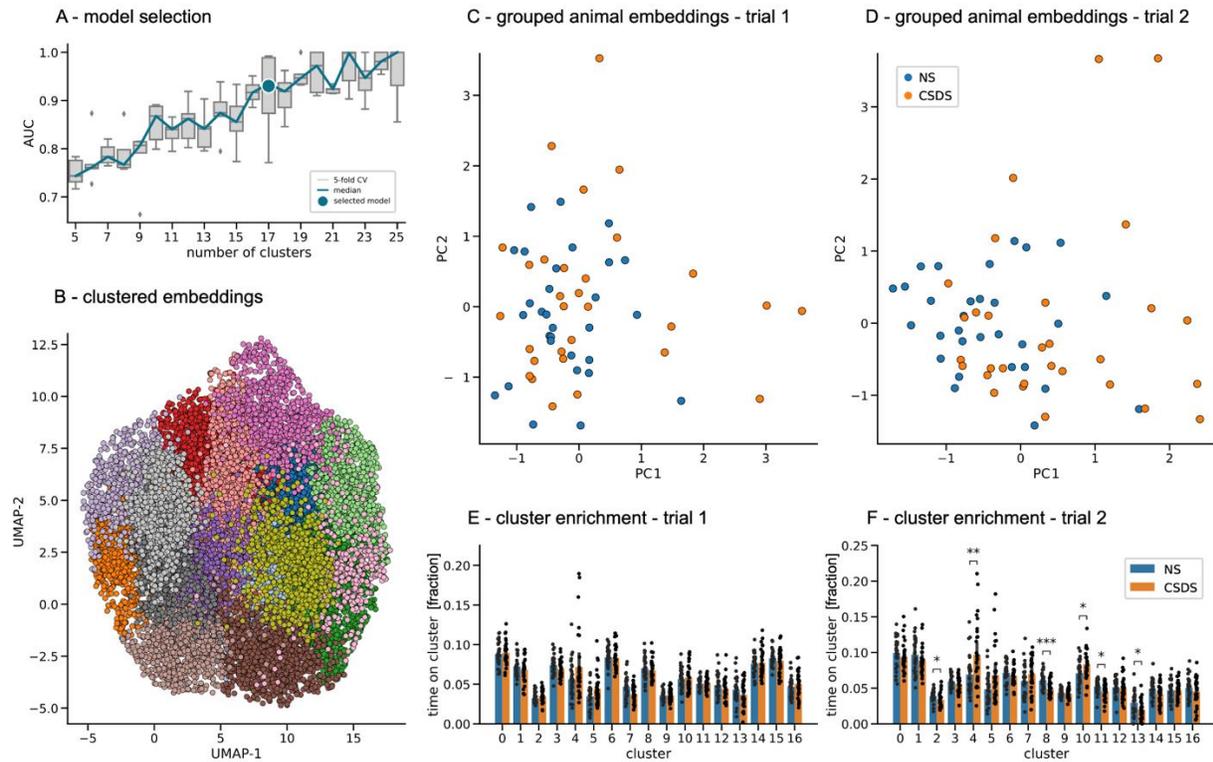
Supplemental Figure 5. Multi-animal unsupervised analyses identify different two-mice behavioral patterns between arenas containing stressed and non-stressed mice during the SI task. A) Cluster selection pipeline results, reporting the area under the ROC curve from a logistic regression classifier discriminating between conditions. A 10-component solution (from a range between 5 and 25) was selected as optimal in a 5-fold (N=5) cross-validation loop (see methods for details). B) Embeddings by time point obtained using DeepOF's unsupervised pipeline. Different colors correspond to different clusters. Dimensionality was further reduced from the original 8-dimensional embeddings using UMAP for visualization purposes. C) Optimal binning of the videos was obtained as the Wasserstein distance between the global animal embeddings of both conditions across a growing window, between the first 10 to 600 seconds for each video at one-second intervals (grey curve). Higher values correspond to larger behavioral differences across conditions. A maximum was observed at 124 seconds, close to the 126 seconds obtained with the single-animal embeddings, and to the stipulated 150 seconds selected based on the SA task literature. The dark green curve depicts the Wasserstein distance across all subsequent non-overlapping bins with optimal length. The decay observed across time is consistent with the hypothesized arousal period in the CSDS cohort, which can be detected also embedding the two-mice system as a whole. D) Representation of the global animal embeddings for the optimally discriminant bin

(124 seconds) per experimental video colored by condition (see methods for details). E) Cluster enrichment per experimental condition (N=26 for NS and N=27 for CSDS) in the first optimal bin (first 124 seconds). Reported statistics correspond to a 2-way Mann-Whitney U non-parametric test corrected for multiple testing using the Benjamini-Hochberg method across both clusters and bins (significant differences observed in clusters 0: $U=1.7e+2$, $p=1.2e-3$, 1: $U=4.9e+2$, $p=8.5e-3$, 3: $U=1.4e+2$, $p=1.4e-4$, 5: $U=8.4e+1$, $p=2.1e-6$, 8: $U=5.3e+2$, $p=1.2e-3$, 9: $U=6.7e+2$, $p=1.4e-8$). Bar graphs represent mean \pm standard deviation of the time proportion spent on each cluster. F) Example heatmap depicting spatial distribution across all experiments (in both conditions) for all clusters. Specific heatmaps for all individual clusters are available in supplemental figure 13). G) Behavioral entropy scores per condition. NS animals show a significantly higher entropy than CSDS animals, which can be attributed to a less predictable exploration of the behavioral space ($U=5.44e+2$, $p=6.15e-4$, N=26 for NS and N=27 for CSDS). Moreover, and in accordance with these results, behavioral entropy shows a significant negative correlation with the presented stress physiology Z-score (supplemental figure 15B). Source data are provided as a Source Data file. Box plots in panels A and G show the median and the inter-quartile range. Whiskers show the full range, excluding outliers as a function of the inter-quartile range.

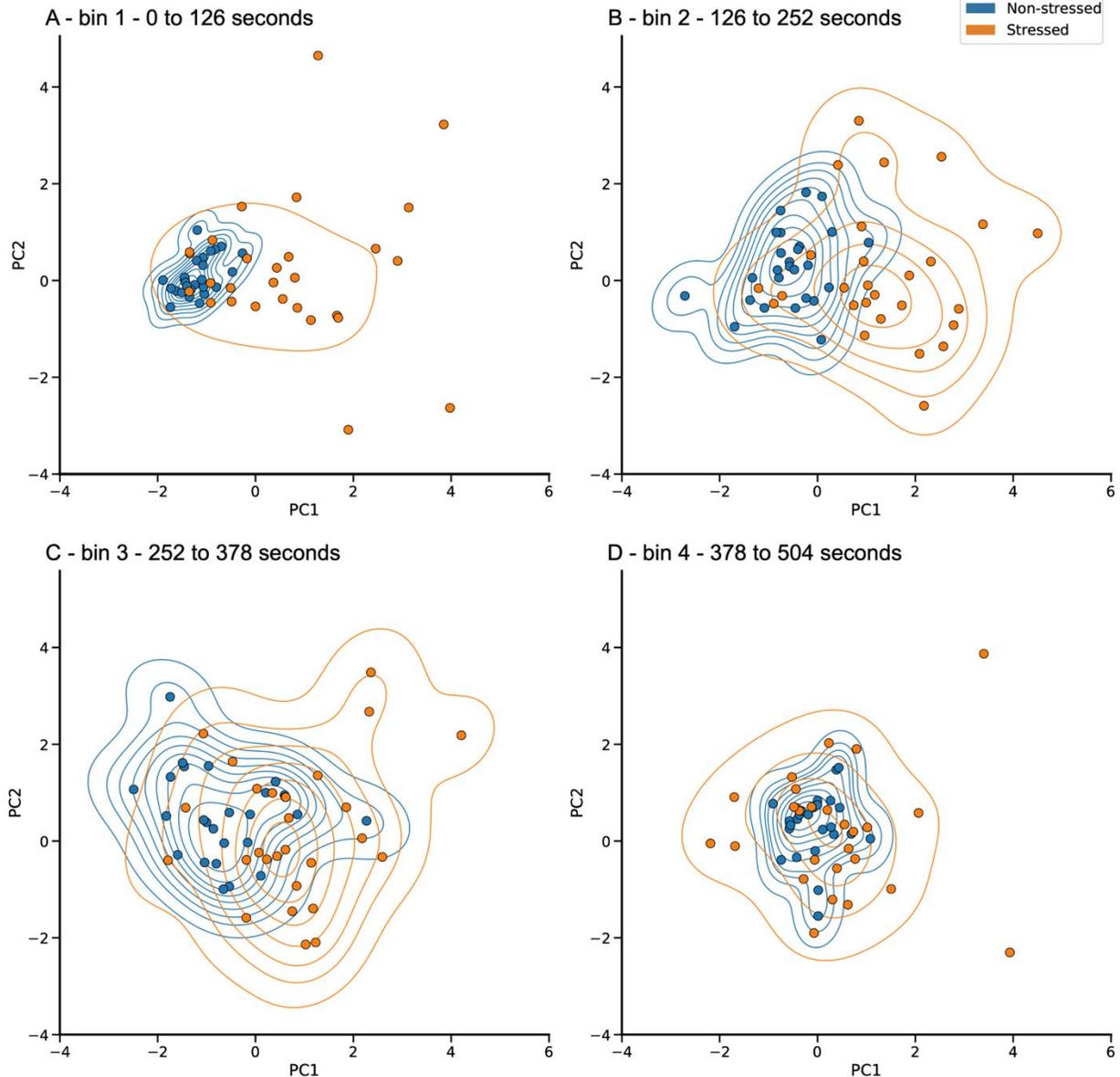


Supplemental Figure 6. Single-animal unsupervised analyses identify different behavioral patterns between stressed and non-stressed mice during the OF task. A) Cluster selection pipeline results, reporting the area under the ROC curve from a logistic regression classifier discriminating between conditions. An 11-component solution (from a range between 5 and 25) was selected as optimal in a 5-fold (N=5) cross-validation loop (see methods for details). B) Embeddings by time point obtained using DeepOF's unsupervised pipeline. Different colors correspond to different clusters. Dimensionality was further reduced from the original 8-dimensional embeddings using UMAP for visualization purposes. C) Optimal binning of the videos was obtained as the Wasserstein distance between the global animal embeddings of both conditions across a growing window, between the first 10 to 600 seconds for each video at one-second intervals (grey curve). Higher values correspond to larger behavioral differences across conditions. A maximum was observed at 595 seconds (green dot), which is consistent with the hypothesized lack of an arousal period in the CSDS cohort in an open field setting with no conspecific. D) Representation of the global animal embeddings for the optimally discriminant bin (595 seconds) per experimental video colored by condition (see methods for details). E) Cluster enrichment per

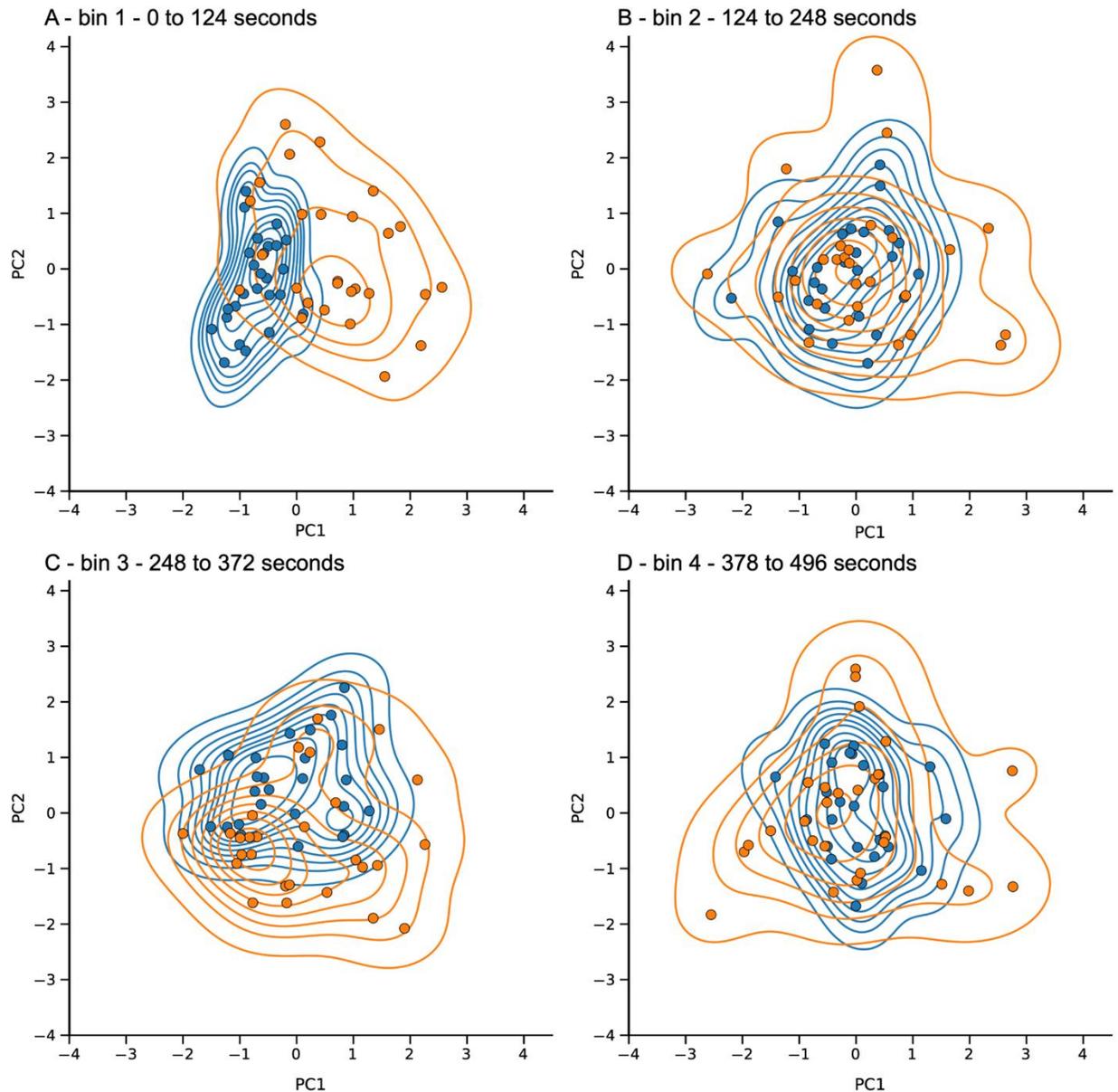
experimental condition (N=26 for NS and N=27 for CSDS) in the first optimal bin (first 595 seconds). Reported statistics correspond to a 2-way Mann-Whitney U non-parametric test corrected for multiple testing using the Benjamini-Hochberg method across both clusters and bins (significant differences observed in clusters 0: $U=2.2e+2$, $p=2.02e-2$, 4: $U=6.1e+2$, $p=5.7e-6$, 5: $U=5.7e+2$, $p=1.3e-4$, 7: $U=5.4e+1$, $p=9.9e-4$, 8: $U=1.8e+2$, $p=2.3e-3$, 9: $U=5.5e+2$, $p=3.7e-4$, and 10: $U=1.5e+2$, $p=2.6e-4$). Bar graphs represent mean \pm standard deviation of the time proportion spent on each cluster. F) Example heatmap depicting spatial distribution across all experiments (in both conditions) for all clusters. Specific heatmaps for all individual clusters are available in supplemental figure 14). G) Behavioral entropy scores per condition. No significant differences are detected between conditions ($U=4.44e+2$, $p=9.98e-2$, N=26 for NS and N=27 for CSDS). Moreover, and in accordance with these results, no significant correlation with the presented stress physiology Z-score was found (supplemental figure 15C). Source data are provided as a Source Data file. Box plots in panels A and G show the median and the inter-quartile range. Whiskers show the full range, excluding outliers as a function of the inter-quartile range.



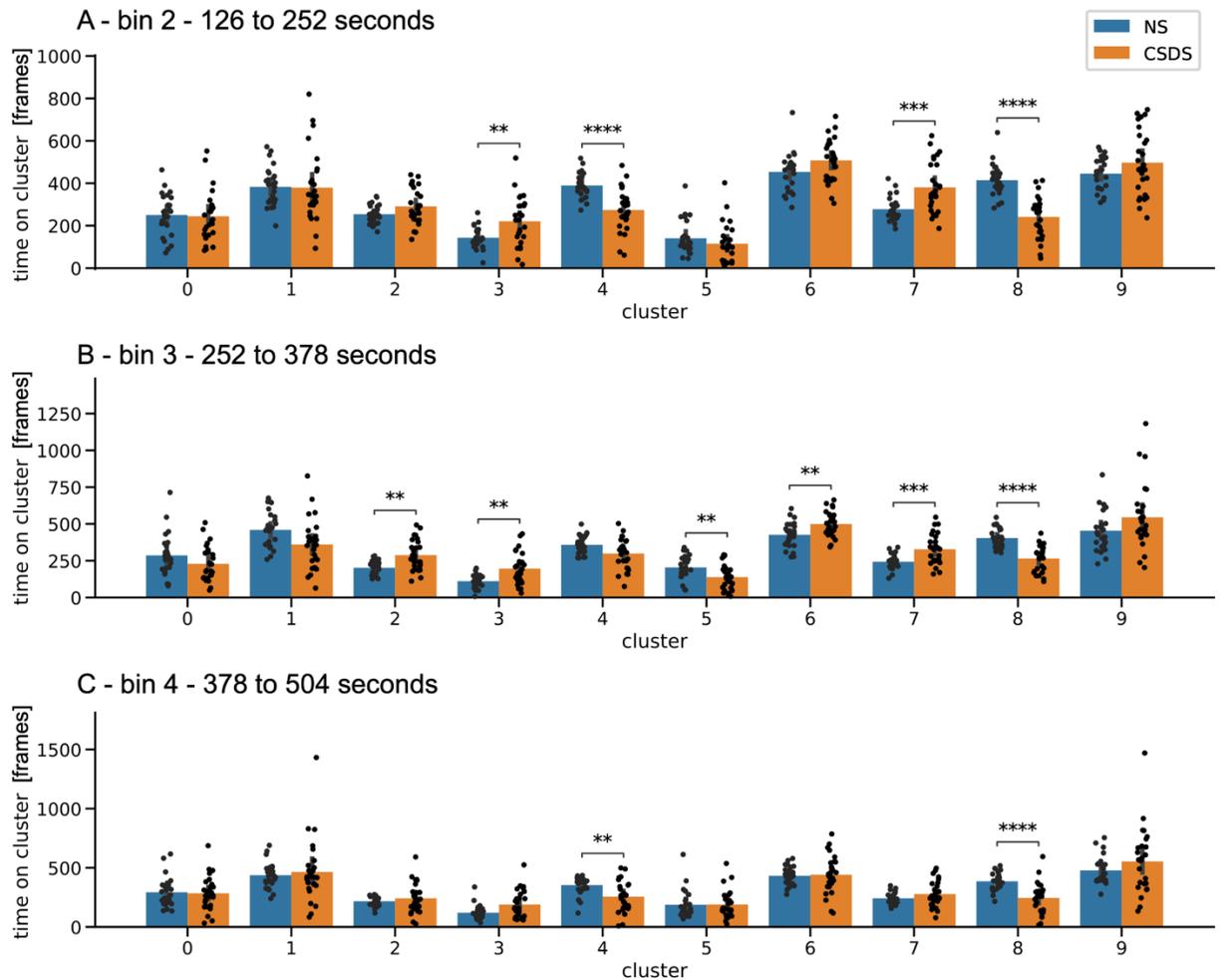
Supplemental Figure 7. Single-animal unsupervised analyses identify mild behavioral differences between stressed and non-stressed mice during the SA task. A) Cluster selection pipeline results. Models ranging from 5 to 25 clusters were trained in a 5-fold (N=5) cross-validation loop using data from both trials together. Area under the ROC curve from a logistic regression classifier discriminating between conditions on the global animal embeddings representing the differential population of each cluster across trials is reported. A 17-component solution was selected as the smallest whose median performance deviated less than one standard deviation from the maximum reached median across all clusters. Boxes in the box plots show the median performance and the inter-quartile range of the data. Whiskers show the full range of the data, excluding outliers as a function of the inter-quartile range. B) Embeddings by time point obtained using DeepOF's unsupervised pipeline. Different colors correspond to different clusters. Dimensionality was further reduced from the original 8-dimensional embeddings using UMAP for visualization purposes. C-D) Representation of the global animal embeddings per experimental video colored by condition, for SA trials one (without conspecific in the cage) and two (with conspecific in the cage). In panel C, as expected, the distributions are further apart. E-F) Cluster enrichment per experimental condition for both SA trials (N=30 for NS and N=30 for CSDS). As expected, trial one shows no significant differences, whereas trial two yields six significantly differentially expressed clusters. Reported statistics correspond to a 2-way Mann-Whitney U non-parametric test corrected for multiple testing using the Benjamini-Hochberg method across both clusters (significant differences for trial two observed in clusters 2: $U=6.1e+2$, $p=1.4e-2$, 4: $U=2.6e+2$, $p=7.3e-6$, 8: $U=7.01e+2$, $p=2.1e-4$, 10: $U=2.8e+2$, $p=1.4e-2$, 11: $U=6.1e+2$, $p=1.7e-2$, and 13: $U=6.1e+2$, $p=1.8e-2$). Bar graphs represent mean \pm standard deviation of the time proportion spent on each cluster. Source data are provided as a Source Data file.



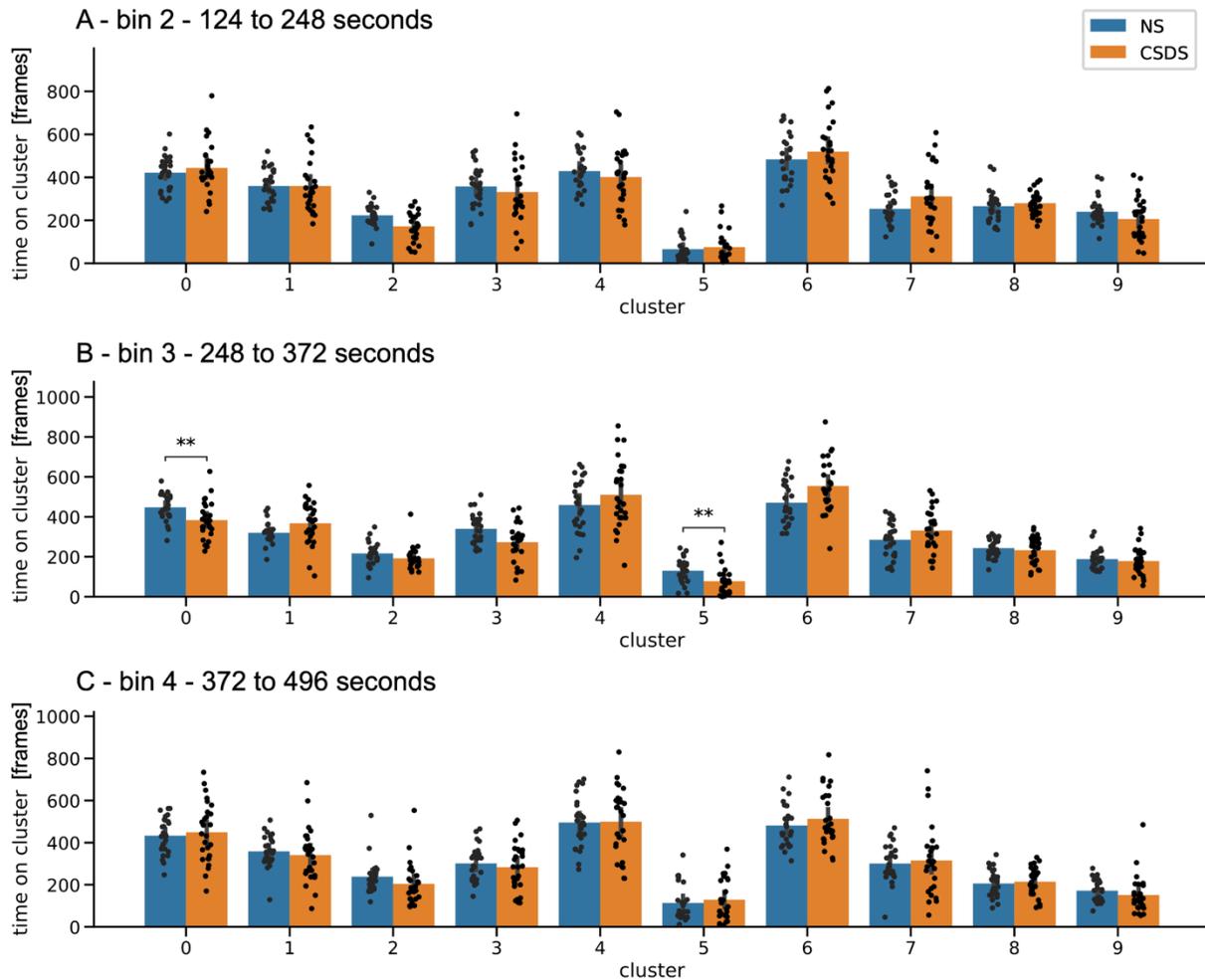
Supplemental figure 8. Global single-animal embeddings across non-overlapping time bins in the SI dataset. A-D) 10-dimensional global single-animal embeddings were obtained as the time proportion spent on each of the 10 clusters in the selected model for the single-animal SI task. Panels A to D show how the distributions matching NS and CSDS animals get closer and closer across non-overlapping consecutive time bins (as quantified using Wasserstein distance in the first four points shown in dark green in figure 6B). The last bin was excluded for visualization purposes. Source data are provided as a Source Data file.



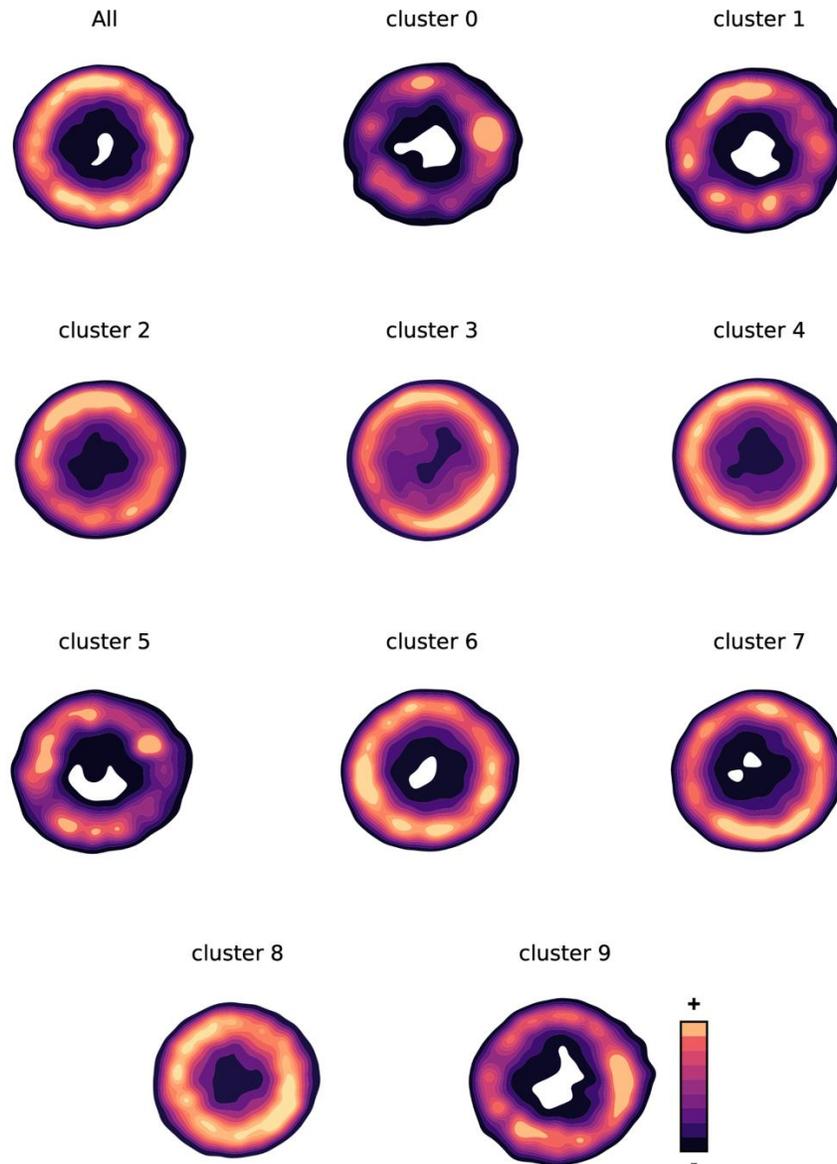
Supplemental figure 9. Global multi-animal embeddings across non-overlapping time bins in the SI dataset. A-D) 10-dimensional global single-animal embeddings were obtained as the time proportion spent on each of the 10 clusters in the selected model for the multi-animal SI task. Panels A to D show how the distributions matching NS and CSDS animals get closer across non-overlapping consecutive time bins (as quantified using Wasserstein distance in the first four points shown in dark green in supplemental figure [9B](#)). The last bin was excluded for visualization purposes. Source data are provided as a Source Data file.



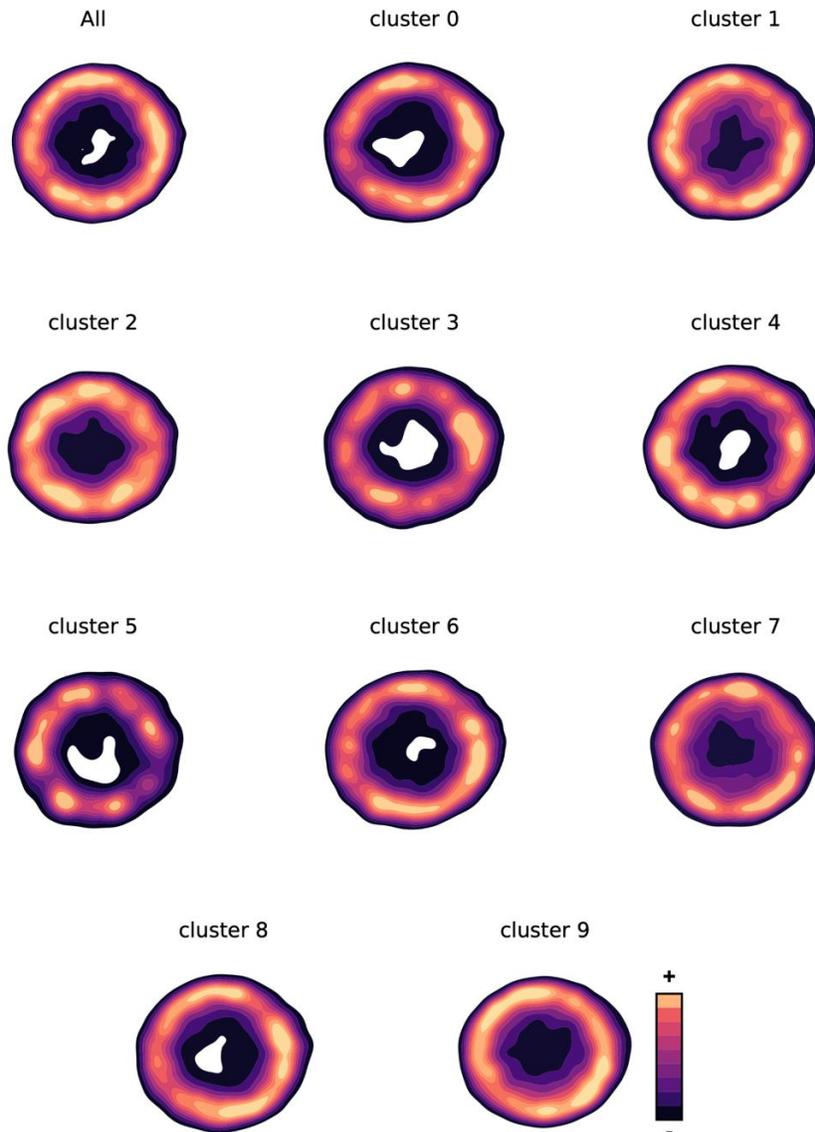
Supplemental figure 10. Cluster enrichment per experimental condition in the second to fourth optimal bins for the single-animal embeddings on the SI task. Reported statistics correspond to a 2-way Mann-Whitney U non-parametric test corrected for multiple testing using the Benjamini-Hochberg method across both clusters and bins. In all cases, N=26 for NS and N=27 for CSDS. A) Second bin (126 to 252 seconds). Significant differences observed in clusters 3: $U=1.9e+2$, $p=6.3e-10$, 4: $U=5.9e+2$, $p=1.4e-5$, 7: $U=1.6e+2$, $p=6.9e-4$, and 8: $U=6.55e+2$, $p=6.3e-8$ B) Third bin (252 to 378 seconds). Significant differences observed in clusters 2: $U=1.8e+2$, $p=1.8e-3$, 3: $U=1.7e+2$, $p=1.2e-3$, 5: $U=4.9e+2$, $p=8.5e-3$, 6: $U=1.9e+2$, $p=7.01e-3$, 7: $U=1.7e+2$, $p=9.6e-4$, and 8: $U=6.3e+2$, $p=6.6e-7$. C) Fourth bin (378 to 504 seconds). Significant differences observed in clusters 4: $U=5.2e+2$, $p=2.5e-5$, and 8: $U=6.02e+2$, $p=6.5e-6$. Bar graphs represent mean \pm standard deviation of the time proportion spent on each cluster. Source data are provided as a Source Data file.



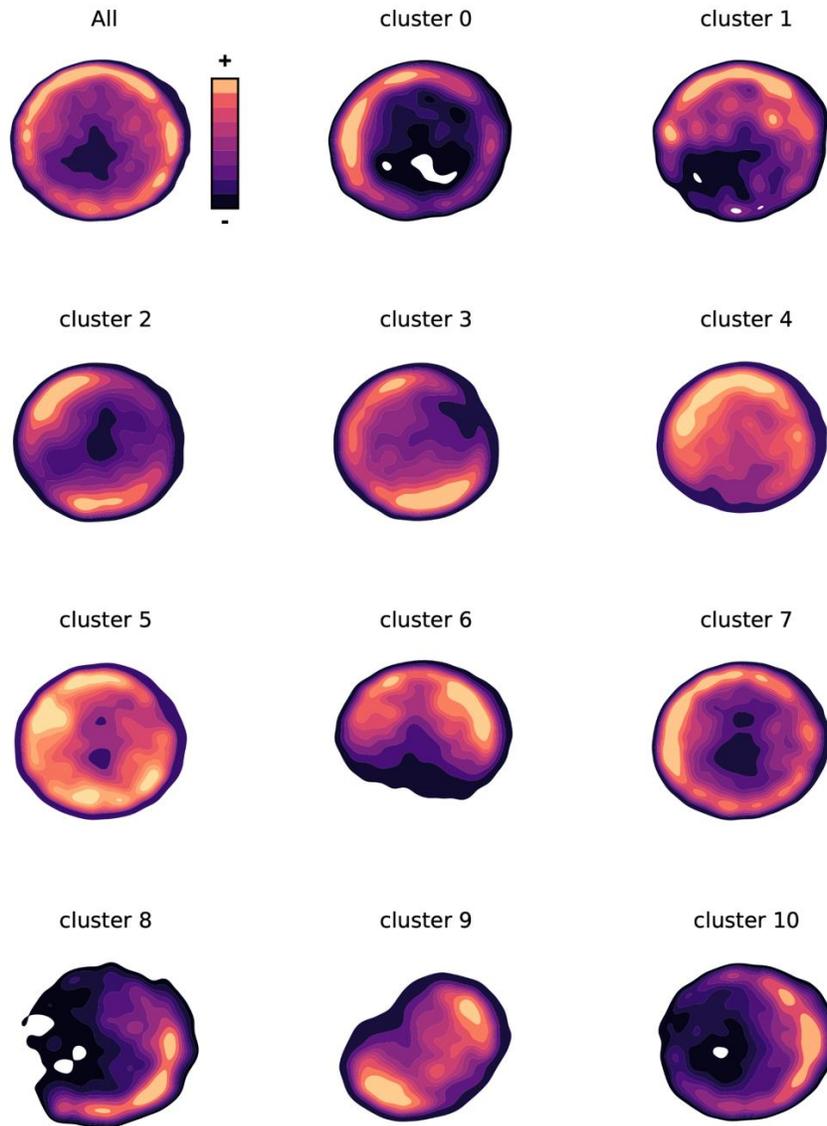
Supplemental figure 11. Cluster enrichment per experimental condition in the second to fourth optimal bins reported for the multi-animal embeddings on the SI task. Reported statistics correspond to a 2-way Mann-Whitney U non-parametric test corrected for multiple testing using the Benjamini-Hochberg method across both clusters and bins. In all cases, N=26 for NS and N=27 for CSDS. A) Second bin (124 to 248 seconds). No significant differences observed. B) Third bin (248 to 372 seconds). Significant differences were observed in clusters 0: $U=5.2e+2$, $p=3.3e-3$, and 5: $U=5.3e+2$, $p=1.6e-3$. C) Fourth bin (372 to 496 seconds). No significant differences were observed. Bar graphs represent mean \pm standard deviation of the time proportion spent on each cluster. Source data are provided as a Source Data file.



Supplemental figure 12. Spatial distribution of clusters obtained using single-animal embeddings in the SI task. Heatmaps include full trajectories of all experiments in both conditions, filtering time points belonging to each obtained cluster, and without filtering (labelled as "all"). White background indicates null population of the area. All clusters enriched in CSDS show lower occupation of the center of the arena than those enriched in NS animals.

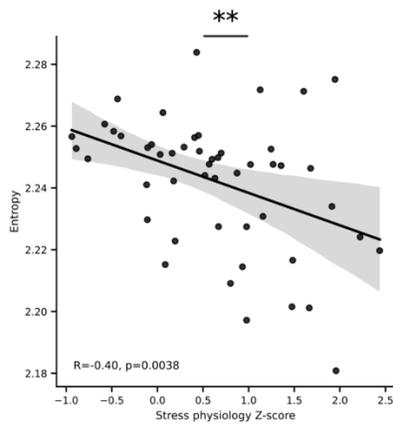


Supplemental figure 13. Spatial distribution of clusters obtained using multi-animal embeddings in the SI task. Heatmaps include full trajectories of all experiments in both conditions, filtering time points belonging to each obtained cluster, and without filtering (labelled as "all"). White background indicates null population of the area. All clusters enriched in CSDS show lower occupation of the center of the arena than those enriched in NS animals.

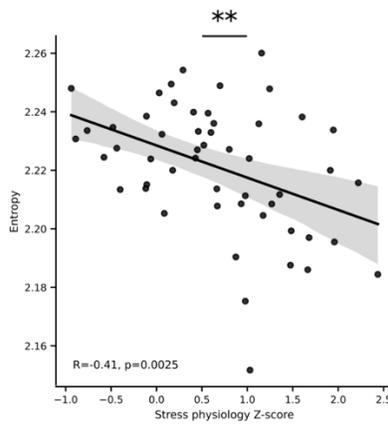


Supplemental figure 14. Spatial distribution of clusters obtained in the OF task. Heatmaps include full trajectories of all experiments in both conditions, filtering time points belonging to each obtained cluster, and without filtering (labelled as "all"). White background indicates null population of the area. All clusters enriched in CSDS show lower occupation of the center of the arena than those enriched in NS animals.

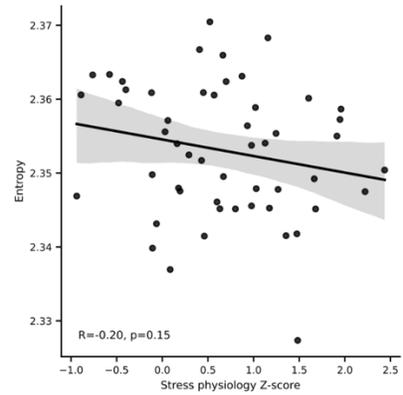
A - SI - single animal



B - SI - multi animal

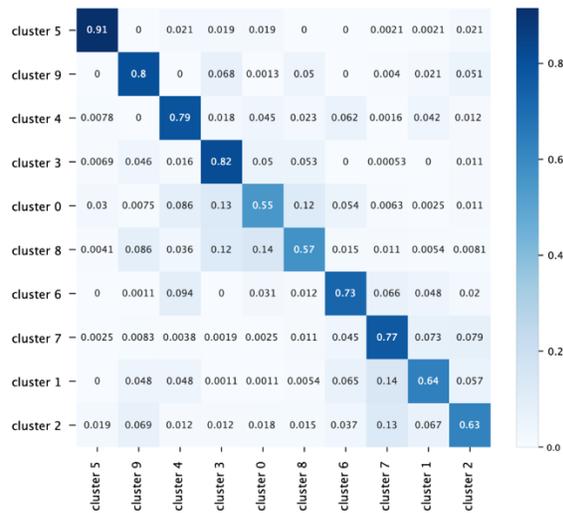


C - OF

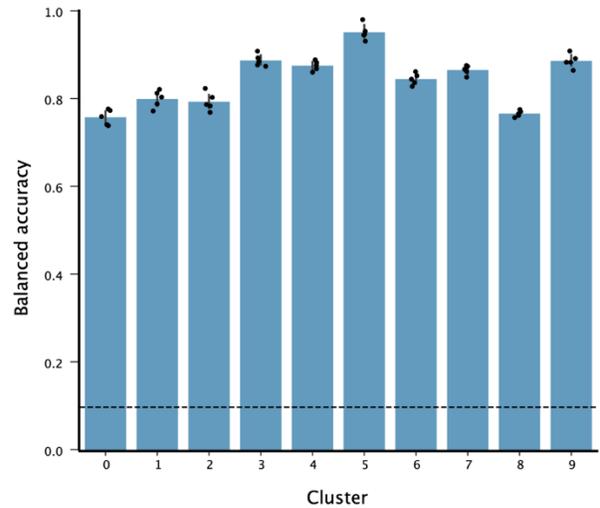


Supplemental Figure 15. Correlation between behavioral entropy and stress physiology Z-score. A) Behavioral entropy of the cluster space obtained with single animal embeddings during the social interaction (SI) task shows a significant negative Pearson correlation with the stress physiology Z-score ($R=-0.40$, $p=3.8e-3$, $N=53$). Error bands represent the 95% confidence band around the mean of the linear model. B) Behavioral entropy of the cluster space obtained with multi-animal embeddings during the social interaction (SI) task shows a significant negative Pearson correlation with the stress physiology Z-score ($R=-0.41$, $p=2.5e-3$, $N=53$). Error bands represent the 95% confidence band around the mean of the linear model. C) Behavioral entropy of the cluster space obtained during the open field (OF) task shows no significant Pearson correlation with the stress physiology Z-score ($R=-0.20$, $p=0.15$, $N=53$). Error bands represent the 95% confidence band around the mean of the linear model. All three tests are two-sided. Source data are provided as a Source Data file.

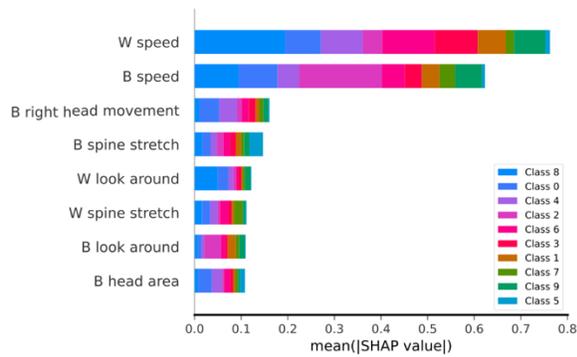
A - cluster detection confusion matrix



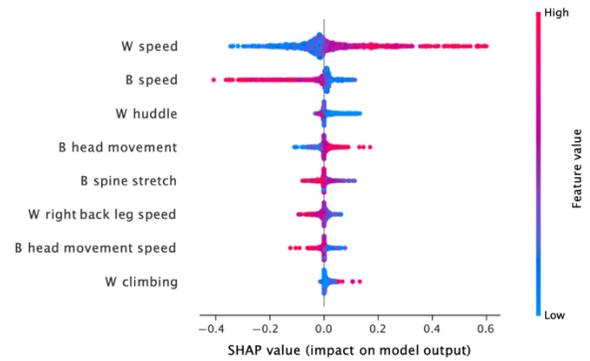
B - cluster detection performance



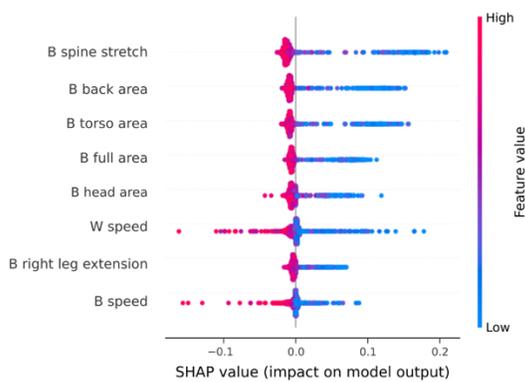
C - SHAP global feature importance



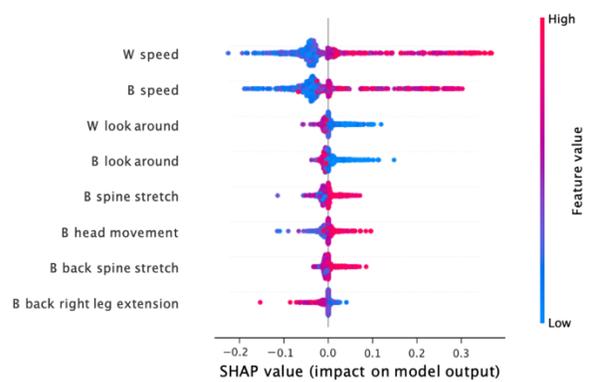
D - SHAP analysis of SI multi-animal cluster 3



E - SHAP analysis of SI multi-animal cluster 5



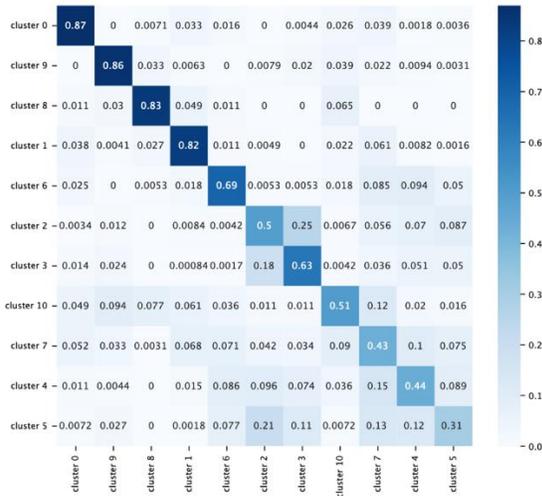
F - SHAP analysis of SI multi-animal cluster 9



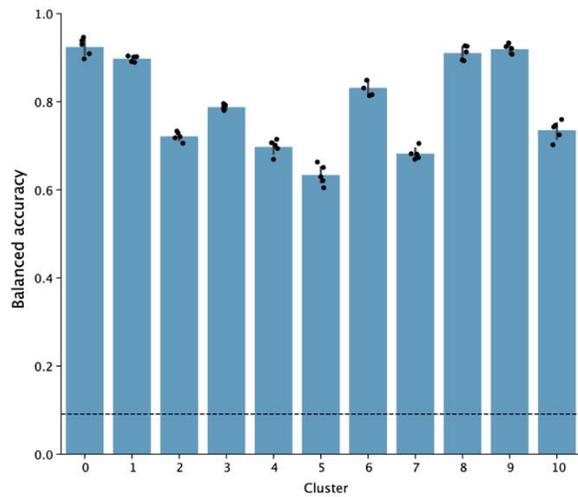
Supplemental Figure 16. SHAP analysis of unsupervised cluster assignments in the multi-animal social interaction task. Gradient boosting machines were trained to map from a predefined set of time series statistics (including body part speeds, distances, distance speeds, areas, area speeds, and supervised annotations for each of the two animals and their interaction) to the previously obtained cluster assignments. A) Confusion matrix obtained from the trained gradient boosting machine classifying

between clusters. Aggregated performance over the validation folds of a 5-fold cross-validation is shown. B) Validation performance per cluster across a 5-fold (N=5) cross-validation loop. Balanced accuracy was used to correct for cluster assignment imbalance. The dashed line marks the expected performance due to chance, considering all outputs. Bars show mean \pm 95% confidence interval. C) Overall feature importance for the multi-output classifier using SHAP. Features in the y-axis are sorted by overall absolute SHAP values across clusters. Classes on the bars are sorted by overall absolute SHAP values across features. D-F) Bee swarm plots for the three most differentially expressed clusters between NS and CSDS mice (3, 5, and 9), identified with the unsupervised DeepOF pipeline on the SI experiments using single-animal embeddings. The depicted plots display the first 8 most important features for each classifier, in terms of the mean absolute value of the SHAP values. Source data are provided as a Source Data file.

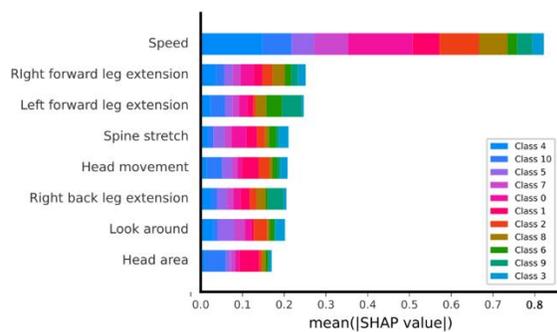
A - cluster detection confusion matrix



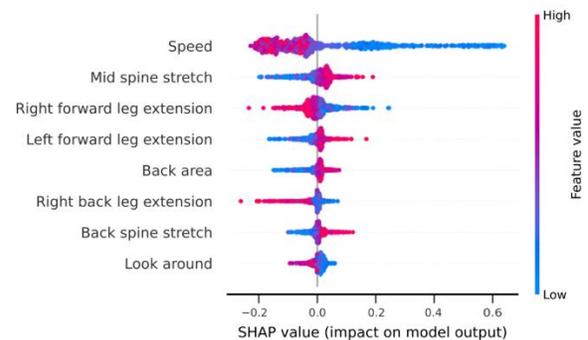
B - cluster detection performance



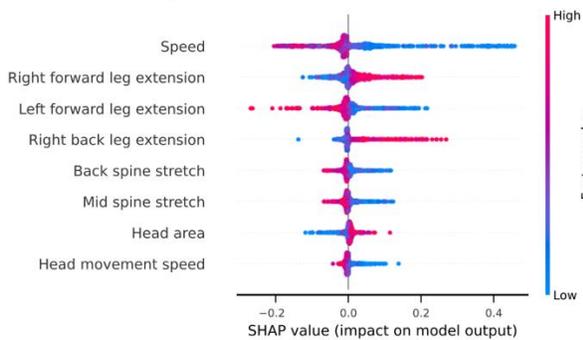
C - SHAP global feature importance



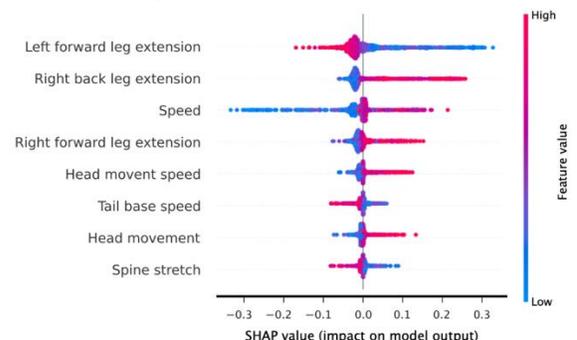
D - SHAP analysis of OF cluster 0



E - SHAP analysis of OF cluster 8



F - SHAP analysis of OF cluster 9



Supplemental Figure 17. SHAP analysis of unsupervised cluster assignments in the open field task.

Gradient boosting machines were trained to map from a predefined set of time series statistics (including body part speeds, distances, distance speeds, areas, area speeds, and supervised annotations) to the previously obtained cluster assignments. A) Confusion matrix obtained from the trained gradient boosting machine classifying between clusters. Aggregated performance over the validation folds of a 5-fold cross-validation is shown. B) Validation performance per cluster across a 5-fold (N=5) cross-validation loop. Balanced accuracy was used to correct for cluster assignment imbalance. The dashed line marks the

expected performance due to chance, considering all outputs. Bars show mean \pm 95% confidence interval. C) Overall feature importance for the multi-output classifier using SHAP. Features in the y-axis are sorted by overall absolute SHAP values across clusters. Classes on the bars are sorted by overall absolute SHAP values across features. D-F) Bee swarm plots for the three most differentially expressed clusters between NS and CSDS mice (4, 9, and 10), identified with the unsupervised DeepOF pipeline on the SI experiments using single-animal embeddings. The depicted plots display the first 8 most important features for each classifier, in terms of the mean absolute value of the SHAP values. Source data are provided as a Source Data file.

4. Automatically annotated motion tracking identifies a distinct social behavioral profile following chronic social defeat stress

Chapter 5

Early life stress affects the acquisition of fear differently between sexes

Bordes, J., Miranda, L., van Doeselaar, L., Brix, L.M., Narayan, S., Yang, H., Mitra, S., Kovarova, V., Springer, M., Müller-Myhsok, B., Schmidt, M.V.

Manuscript in preparation

Early life stress affects the acquisition of fear differently between sexes

Joeri Bordes¹, Lucas Miranda^{2,3}, Lotte van Doeselaar^{1,3}, Lea Maria Brix^{1,3}, Sowmya Narayan^{1,3}, Huanqing Yang¹, Shiladitya Mitra¹, Veronika Kovarova^{1,3}, Margherita Springer¹, Bertram Müller-Myhsok², Mathias V. Schmidt¹

Affiliations:

¹Research Group Neurobiology of Stress Resilience, Max Planck Institute of Psychiatry, 80804 Munich, Germany

²Research Group Statistical Genetics, Max Planck Institute of Psychiatry, 80804 Munich, Germany

³International Max Planck Research School for Translational Psychiatry (IMPRS-TP), 80804 Munich, Germany

Abstract

Exposure to early life stress (ELS) has a detrimental effect on both the physiological and behavioral outcomes and ultimately can lead to increased vulnerability toward stress-related disorders, such as PTSD. The prevalence of PTSD is strongly influenced by sex with around twice as many women suffering from PTSD compared to men, which is partially mediated via genetic risk variants. Understanding the role of sex on the interaction of PTSD risk genes, such as *FKBP5*, with behavioral consequences of trauma exposure is crucial to further unraveling the underlying neurobiological pathways of PTSD. The development of unsupervised behavioral analysis tools has allowed for in-depth behavioral analysis that can explore previously unknown behavioral patterns. This is crucial to increase the understanding of the behavioral outcome related to stress-induced fear memory formation. The current study investigates the sex-specific effects of ELS by using the limited nesting and bedding (LBN) stress paradigm on *Fkbp5* expression in the basolateral amygdala (BLA) and hippocampus (HIP) while exploring the fear memory formation using unsupervised clustering analysis. A sex-specific effect of LBN exposure was observed for a dysregulated hypothalamic-pituitary-adrenal (HPA) axis, in which baseline corticosterone (CORT) was increased in females directly after stress exposure at postnatal day 9, while adrenal weight and *Fkbp5* expression in the CA1 HIP region were increased in ELS males in adulthood. Moreover, specific aspects of fear-related behavior were altered by LBN in adulthood, including the passive fear response via freezing behavior in both the acquisition and the retrieval of fear. In addition, an unsupervised deep phenotyping strategy was implemented to understand the different behaviors expressed during fear acquisition. The fear acquisition in LBN females showed a higher active fear response, which was not observed in LBN males. Therefore, the in-depth behavioral analysis using unsupervised clustering provides an additional layer to explore the fear-related behaviors without prior assumptions and therefore allows for hypothesis-generating behavioral analysis, which ultimately can lead to a better understanding of the stress-induced behavioral phenotype.

Introduction

Early life stress (ELS) exposure, such as child abuse or neglect, has severe long-lasting negative behavioral and physiological consequences in adulthood, including, among others, an altered neuroendocrine function¹⁻³ and morphological changes in the brain^{4,5}. This ultimately leads to an increased risk and persistence of stress-related disorders, such as post-traumatic stress disorder (PTSD)⁶⁻⁸. Human genetic studies have identified that stress-related disorders are partially mediated by different genomic variations⁹⁻¹⁵. In particular, the combination of childhood trauma and specific genetic risk variants of the *FKBP5* gene, which encodes for the glucocorticoid receptor (GR) co-chaperone FK506 binding protein 51 (FKBP51), has been found to increase the risk for developing PTSD¹⁶⁻¹⁹. FKBP51 stabilizes the GR-complex structure, thereby decreasing the binding to glucocorticoids and hampering the nuclear translocation of the GR complex. Altered FKBP51 expression has been implicated in dysregulation of the hypothalamic-pituitary-adrenal (HPA) axis²⁰, which in turn increases the risk for psychopathology^{21,22}. An additional important factor influencing the susceptibility and severity of PTSD is sex. The prevalence of PTSD is around twice as high in women (11.0%) compared to men (5.4%)²³. However, the increased risk for PTSD development in women could not be entirely explained by differences in the event type or severity of the traumatic event, which might indicate a potential underlying biological mechanism^{24,25}. Therefore, investigating the underlying mechanisms related to ELS exposure in a sex-specific way is crucial to advance the understanding of the neurobiological mechanism of PTSD.

An increasing number of clinical studies are pointing to corticolimbic structures as strongly affected in size and functioning by ELS exposure, in particular the basolateral amygdala (BLA) and hippocampus (HIP)²⁶⁻³⁰. Moreover, the BLA and HIP continue their development in function and morphology during the early postnatal period³¹⁻³³, rendering this developmental period especially vulnerable to environmental insults. In addition, both BLA and HIP display a high expression of the GR and mineralocorticoid receptor (MR)³⁴⁻³⁷, which makes them particularly vulnerable to ELS exposure. Moreover, animal models have identified a particularly high expression of the *Fkbp5* gene in the BLA and HIP and observed a stress-dependent upregulation of *Fkbp5* expression in those regions³⁸.

Alterations of anxiety and fear behavior are a central hallmark of a PTSD-like phenotype in animal models. A study using a mouse model overexpressing the human *FKBP5* gene in the forebrain showed that elevated FKBP5 expression in combination with ELS exposure increases anxiety-related behavior, which was more pronounced in females³⁹. There is an increasing body of evidence that ELS exposure affects rodents in a sex-specific manner⁴⁰⁻⁴⁴. However, the effects of fear memory formation on ELS have not

been explored in detail. Exposure to fear conditioning in rodents has been shown to strongly activate the BLA and HIP in a time-dependent manner ⁴⁵. Therefore, the important role of the BLA and HIP in ELS exposure indicates that the formation of fear acquisition and memory could be an important target for ELS exposure. Previous research has shown that ELS reduces fear expression during contextual as well as auditory fear memory retrieval in males, which is linked to a reduction of synaptic plasticity markers in the dorsal HIP ⁴⁶, but the sex-dependent effects and exact behavioral mechanism remain to be uncovered.

In the current study, we investigate the sex-specific effects of ELS exposure by using the established limited bedding and nesting (LBN) paradigm on *Fkbp5* expression and fear acquisition and retrieval. Utilizing an unsupervised deep phenotyping strategy, we can show that specific aspects of fear behavior and memory are altered by ELS in a sex-specific manner. These behavioral alterations align with *Fkbp5* expression changes in BLA and HIP.

Materials and methods

Animals

Adult male and female C57/Bl6N mice (age between 2-3 months of age) were obtained from the in-house facility of the Max Planck Institute of Psychiatry and used for breeding (F_0). Animals from the F_1 generation were used as experimental animals and were weaned at P25 in groups of maximum four animals with their littermates. Animals were housed in individually-ventilated cages (IVC; 30cm×16cm×16cm connected by a central airflow system: Tecniplast, IVC Green Line—GM500). All animals were kept under standard housing conditions; 12h/12h light-dark cycle (lights on at 7 a.m.), temperature $23\pm 1^\circ\text{C}$, humidity 55%. Food (Altromin 1324, Altromin GmbH, Germany) and tap water were available ad libitum. All experimental procedures were approved by the committee for the Care and Use of Laboratory Animals of the government of Upper Bavaria, Germany. All experiments were in accordance with the European Communities Council Directive 2010/63/EU.

ELS paradigm: limited bedding and nesting

ELS was performed using the LBN paradigm to induce chronic stress towards the mother and pups during P02 to P09, as previously described by Rice et al.⁴⁷. At P02, all litters were transferred to new IVCs and randomly assigned to the stressed or nonstressed condition. If necessary, the litters were culled to a maximum of 10 animals per litter. The stressed litters were placed on a stainless-steel mesh (McNichols) and were provided with limited nesting material (1/2 square of Nestlets, Indulab). The nonstressed animals were placed in an IVC with a standard amount of bedding material and were provided with a sufficient amount of nesting material (2 squares of Nestlets). All litters were left undisturbed until P09, after which they returned to standard housing conditions. The pups were weaned in same-sex groups with a maximum of four animals per cage.

Adult behavioral testing

At 3 months of age, a cohort of both males and females were tested on a fear conditioning protocol, containing fear acquisition (day 1), and the subsequent recall of contextual fear memory (day 2) and auditory fear memory (day 3). The behavioral tests were performed between 8 a.m. and 11 a.m. in the same room as the housing facility.

Fear conditioning

The fear conditioning protocol was performed as previously described ⁴⁸. Data were recorded and analyzed using the ANY-maze 7.2 software (Stoelting, Ireland), in which the percentage of the time freezing was calculated. Furthermore, the fear acquisition data were subsequently analyzed using DeepLabCut version 2.2b7 ⁴⁹ and DeepOF version 0.1.6 ⁵⁰ for the unsupervised analysis pipeline.

Fear acquisition

The fear acquisition consisted of placing the mice into a cube-shaped fear conditioning chamber (Bioseb, France) with a metal grid floor to provide electric shocks. At the start of the test, the chamber light was switched on, and after an initial habituation time of 3 min, the mice were exposed to five conditioned-unconditioned stimulus pairings (auditory conditioning stimulus: 30 sec, 9kHz, 80dB tone & unconditioned stimulus 0.5 sec, 0.6mA foot shock) with an inter-trial interval (ITI) of 5 mins. 1 min after the last foot shock the animals were returned to their home cage. Before and after each trial, the conditioning chamber was thoroughly cleaned with 70% EtOH. The calculation of the mean freezing statistics was performed using the average of tones 2-5, leaving out the first tone, as no shock history was present at that moment. The mean ITI freezing was calculated using all four ITIs.

Contextual fear memory

Contextual fear memory was performed 24 hours after initial fear acquisition. The same setup was used as by fear acquisition, except that no conditioned-unconditioned stimulus protocol was executed. The test endured for a total of 5min in which only the chamber light was switched on, and again before and after each trial, the conditioning chamber was thoroughly cleaned with 70% EtOH.

Auditory fear memory

The consolidation of auditory fear memory was performed two days after the fear acquisition. The set-up was replaced by a novel and neutral context, which differed in material (plexiglass), shape (circular), and surface texture, as no grid was present at the bottom of the set-up. In addition, the cleaning solution was changed in odor, using 1% acetic acid. This allowed for measuring the fear response specifically towards the tones, without the interference of the context. The chamber light was switched on at the start of the test, after which mice were left undisturbed for an initial 1 min habituation phase. Then, the mice were exposed to the same tones as heard in the fear acquisition (30 sec, 9kHz, 80dB) for 15 times with a 1.5 min ITI. 1 min after the last tone the animals were returned to their home cage. The mean tone and ITI freezing were calculated using all tones and ITIs.

Unsupervised analysis of the fear conditioning data

An additional unsupervised analysis was performed for the fear acquisition data in both males and females, in order to obtain a more in-depth analysis of the behavioral differences between conditions and sexes during the ITIs and tones 2-5. First, pose estimation was performed on the raw data videos using DeepLabCut version 2.2b7 (single animal mode). DeepLabCut pose estimation analysis was performed using 11 body parts, including the nose, left and right ears, three points along the spine (including the center of the animal), all four extremities, and the tail base.

Subsequently, DeepLabCut annotated datasets were processed and analyzed using DeepOF v0.2, as described previously⁵⁰. In brief, DeepOF preprocesses the DeepLabCut annotated data by performing alignment and centering of the coordinates, calculating the distances between body parts, the angles, and areas of specific regions of each available body part, as well as their speeds, accelerations, and higher-order derivatives. The unsupervised analysis of the fear acquisition data was performed on the entire video length and utilized the same model for the male and female data, in order to make cluster interpretation between sexes possible. The interpretation of the clusters was explored by visual inspection of representative video snippets for each specific cluster. Representatives were selected as instances with an assigned cluster assignment confidence greater or equal than 0.9. In addition, Shapley additive explanations (SHAP) were utilized to rank feature importance per cluster, in order to further understand the expressed behavior within clusters.

Physiological measurements

One week following the behavioral tests, adult animals were weighed and subsequently sacrificed by decapitation, after which trunk blood was collected in EDTA-coated microcentrifuge tubes (Kabe Labortechnik, Germany) and directly transferred to ice. Samples were centrifuged at 4°C for 15min at 8.000 rpm, after which plasma was removed and kept transferred for storage at -80°C. A separate cohort of mice was used to obtain corticosterone (CORT) measurements directly after the stress at P09. On the morning of P09, litters were kept in their cage, while first the mother was sacrificed, and then subsequently the pups were sacrificed, keeping them in their nest as long as possible to minimize the influence of acute stress exposure. Trunk blood was collected and processed as described for adult blood samples. Plasma CORT levels were measured in duplicates using radioimmunoassay following the manufacturer's protocol (MP Biomedicals, Eschwege, Germany). Adrenals were dissected and kept at 4°C in saline (0.9% NaCl) until further processing, which included the removal of all surrounding fat tissue and

weighing. The relative adrenal weight was calculated by dividing the total body weight before sacrifice by the total adrenal weight, including the adrenals from both sides.

In-situ hybridization of *Fkbp5*

The *Fkbp5* mRNA profile was determined using radio-active *in-situ* hybridization labeling as described previously³⁸. In brief, the animals were either sacrificed directly after the stress exposure at P09, or in adulthood at 2 months of age. After decapitation, the brains were removed and snap-frozen using 2-methyl butane (kept on dry ice) and stored at -80°C until further use. Brains were sliced using a cryostat in 20 µm sagittal sections, which resulted in a series of the BLA and dorsal HIP slides that were thaw-mounted on Super Frost Plus Slides and stored at -20°C. The *in-situ* hybridization sections were removed from -20°C, left to dry at room temperature, fixated with 4% paraformaldehyde, and subsequently dehydrated using a series of increasing concentrations of ethanol. Then, the hybridization buffer was equally spread out over the different slides containing the radioactive ³⁵S-UTP-labeled *Fkbp5* riboprobe and incubated overnight at 55°C. On the next day, the sections were rinsed, incubated with RNase A, desalted, and dehydrated, after which the radioactive slides were exposed to Kodak Biomax MR films (Eastman Kodak Co., Rochester, NY) and developed after an exposure time of 12 days. Films were digitized and the regions of interest were identified using the mouse brain atlas (<https://developingmouse.brain-map.org/static/atlas>). The expression was determined by optical densitometry with the ImageJ software (NIH, Bethesda, MD, USA). The expression was averaged per brain region per animal and subtracted by the background signal of a nearby structure that did not express the *Fkbp5* gene. A distinction was made between important subregions of the dorsal HIP, in which separate measurements were obtained for the CA1, CA2-3, and the dentate gyrus (DG).

Statistics

Statistical analyses and graphs were made using RStudio (with R 4.1.1), except for the unsupervised DeepOF analysis, which was performed using Python (v 3.9.13). Different batches of animals were used for the adult fear conditioning behavior (female: LBN n=10, NS n=10 & male: LBN n=11, NS n=11) and the in-situ hybridization experiments (female P09: LBN n=4, NS n=6 & male P09: LBN n=5, NS n=3, female adult (2 months): LBN n=5, NS n=4 & male adult (2 months): LBN n=3, NS n=4). All animals were used for statistical analysis unless stated otherwise. During the contextual fear memory, 2 nonstressed male animals were excluded from the analysis due to technical difficulties. Data were tested for the corresponding statistical assumptions, which included the Shapiro-Wilk test for normality and Levene's test for heteroscedasticity. If assumptions were violated the data were analyzed using non-parametric variants of the test. The group comparisons were analyzed using the independent samples t-test (T) as a parametric option, Welch's test (We), if data was normalized but heteroscedastic, or the Wilcoxon test (Wx) as a non-parametric option. The time-binned data (fear conditioning) was analyzed using the two-way repeated measures ANOVA with the phase (e.g. tones) as a within-subject factor and the condition (nonstressed vs. stressed) as a between-subject factor. Data that showed a significant main effect were further analyzed with the post-hoc Bonferroni test (parametric) or the Kruskal-Wallis test (non-parametric). P-values were adjusted for multiple testing using the Bonferroni method. The timeline and bar graphs are presented as mean \pm standard error of the mean (SEM). Data were considered significant at $p < 0.05$ (*), and further significance was represented as $p < 0.01$ (**), $p < 0.001$ (***), and $p < 0.0001$ (****).

Results

The physiological hallmarks of LBN

Chronic stress during early life was induced to assess the sex-dependent stress effects directly after early stress exposure and in adult age (figure 1A). A common phenomenon of LBN exposure is the reduction in body weight at P09⁴⁷, therefore chronic stress exposure was successfully induced in the current study for both sexes, as a significant decrease in body weight at P09 was found (figure 1B). Interestingly, the stress-induced reduction in body weight was sustained in female adult age, but not in males (figure 1C). A marker for chronic stress exposure and dysregulation of the HPA axis is the relative weight of the adrenals, which in adult age was not altered in females, but was significantly elevated in stressed males (figure 1D). In addition, as a proxy for stress exposure, the CORT levels were obtained directly after the chronic stress exposure at P09, the female LBN animals showed a significant elevation of CORT levels, whereas in males no elevated cort levels were observed (figure 1E). At adult age, no difference in basal CORT levels was observed for stress exposure in both sexes (figure 1F).

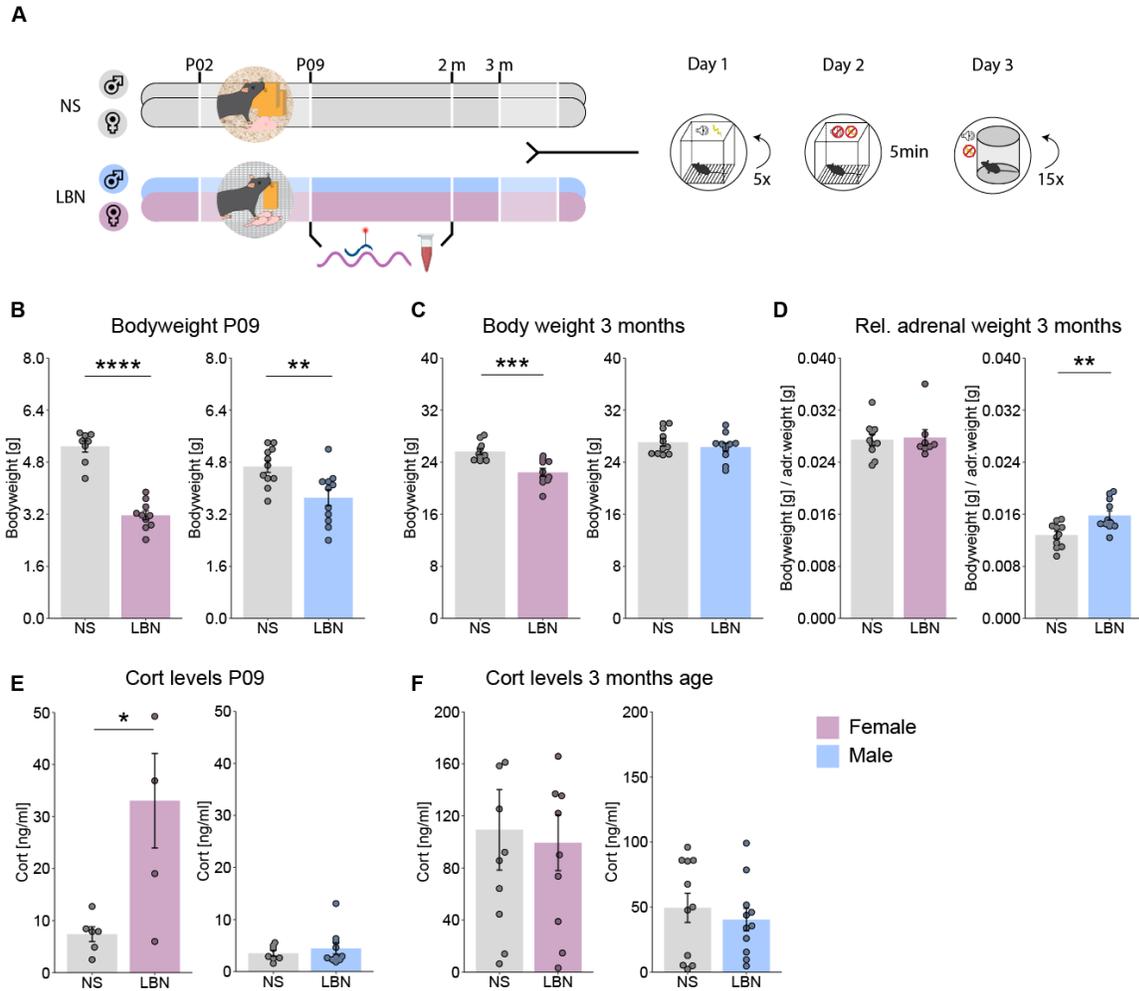


Figure 1. Physiological stress hallmarks of LBN. **A)** Experimental timeline for LBN paradigm and behavioral testing. **B)** Significant decrease in body weight was observed after LBN exposure at P09 for females ($W_x=80$, $p<0.0001$), and males ($T(20)=3.13$, $p=0.005$). **C)** During adult age (3 months) the body weight was significantly reduced in females ($T(18)=4.16$, $p=0.0006$), but not in males ($T(20)=0.88$, $p=0.39$). **D)** The relative adrenal weight was not altered in females ($W_x=45$, $p=0.70$), but was significantly increased in stressed males ($T(20)=-3.3$, $p=0.003$). **E)** At P09, CORT levels were significantly elevated in stressed females ($W_x=4$, $p=0.05$), whereas this was not the case for males ($W_x=31$, $p=0.74$). **F)** At adult age, the CORT levels were not altered by stress for both females ($T(18)=0.26$, $p=0.79$) and males ($T(20)=0.64$, $p=0.53$).

Fkbp5 mRNA levels in the BLA and dorsal hippocampus

Fkbp5 mRNA expression was assessed directly after LBN exposure at P09 and in adult age in the BLA (figure 2A) and dorsal HIP, separating the important subregions; CA1, CA2-3, and DG (figure 2B). Female and male data at P09 did not show a stress-induced difference of *Fkbp5* expression in the BLA, CA1, CA2-3, and DG (figure 2C, D), however, an indication of elevated *Fkbp5* expression in stressed animals could be observed in males, but this was not significant (figure 2D). Moreover, also at adult age, the female data did not indicate any *Fkbp5* expression differences between stress conditions (figure 2E). However, the adult male observations showed a significant stress-induced increase of *Fkbp5* expression in the CA1 region, but not in the BLA, CA2-3, and DG regions (figure 2F). Moreover, both nonstressed and stressed showed an age-dependent expression pattern of *Fkbp5* regardless of sex in the dorsal HIP. At P09 *Fkbp5* expression is the highest in the CA2-3 and similar in the CA1 and DG, whereas at p56, the *Fkbp5* expression was the highest in the DG, then the CA2-3, and then the CA1 (figure 2C-F).

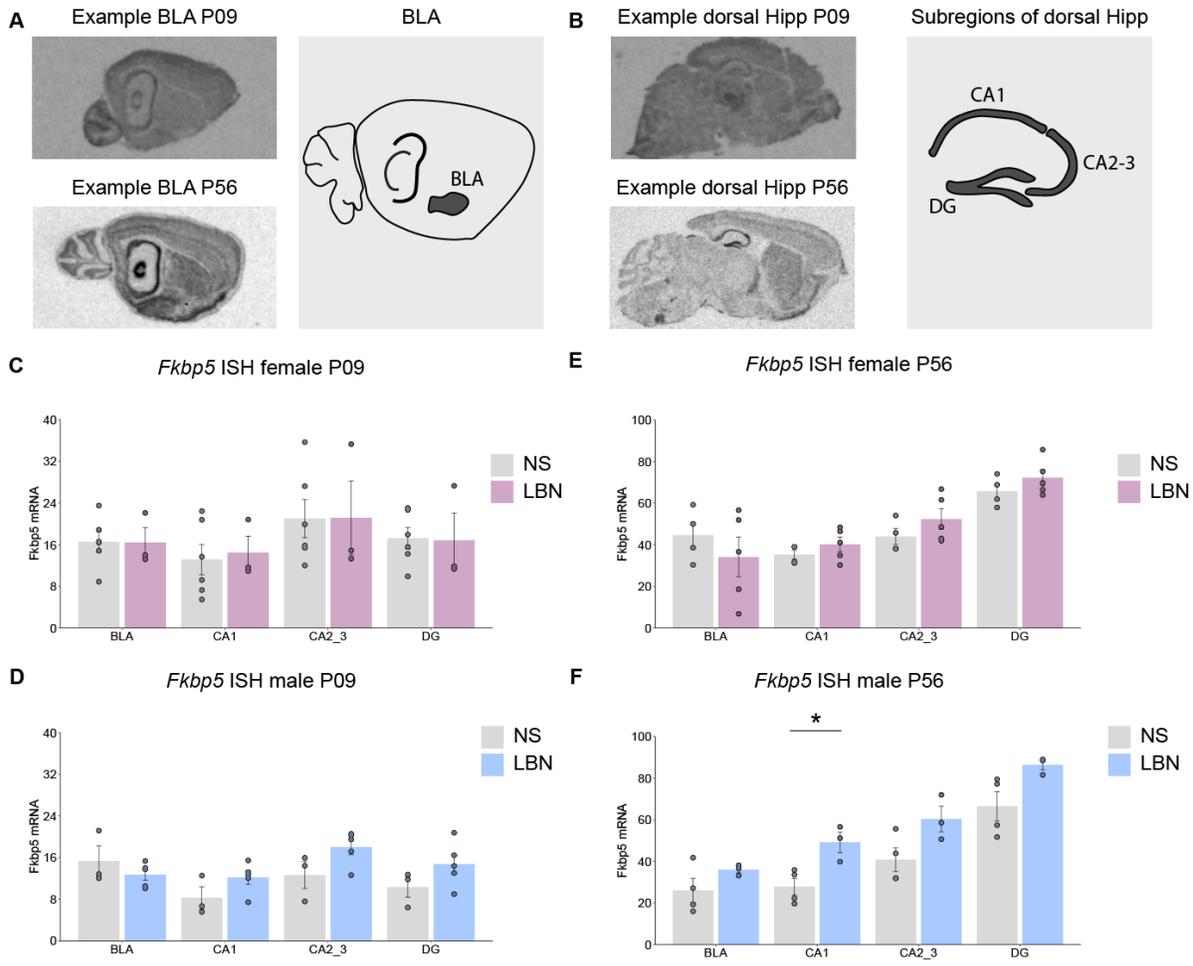


Figure 2. In-situ hybridization of *Fkbp5* mRNA in the BLA and dorsal hippocampal subregions. A) In-situ hybridization scan of *Fkbp5* mRNA expression at P09 and P56 in the BLA. **B)** In-situ hybridization scan of *Fkbp5* mRNA expression at P09 and P56 in the subregions of the dorsal HIP. **C)** No differences were observed in females at P09 in the BLA ($T(7)=0.03$, $p=0.98$), CA1 ($T(7)=-0.28$, $p=0.79$), CA2-3 ($Wx=11$, $p=0.71$), and DG ($T(7)=0.09$, $p=0.93$). **D)** No differences were observed in males at P09 in the BLA ($T(6)=1.03$, $p=0.34$), CA1 ($T(6)=-1.66$, $p=0.15$), CA2-3 ($T(6)=-1.98$, $p=0.096$), and DG ($T(6)=-1.49$, $p=0.19$). **E)** No differences were observed in females at P56 in the BLA ($T(7)=0.86$, $p=0.42$), CA1 ($T(7)=-1.14$, $p=0.29$), CA2-3 ($T(7)=-1.27$, $p=0.25$), and DG ($T(7)=-1.21$, $p=0.27$). **F)** A significant difference for elevated *Fkbp5* mRNA expression was observed in the stressed condition for the males at P56 in the CA1 ($T(5)=-3.38$, $p=0.020$), but not in the BLA ($T(5)=-1.44$, $p=0.21$), CA2-3 ($T(5)=-2.30$, $p=0.070$), and DG ($We(3.67)=-2.69$, $p=0.056$).

Freezing behavior is affected by LBN exposure in a sex-specific manner

During the acquisition of fear conditioning, the typical increase of freezing behavior over the different tone representations was observed in both females and males, regardless of the stress condition (figure 3A-D). However, an overall decrease during the fear acquisition in freezing behavior was observed during the tone representations in males, but not females (figure 3B, D). Moreover, the exploration of the ITIs during the fear acquisition phase showed that the freezing during the individual ITIs was not significantly altered in females based on the stress condition (figure 3E), but there was a significant reduction in the freezing response of stressed females in the overall mean ITIs (figure 3F), which was not observed in males (figure 3G-H). When looking at the recall of fear, the mean freezing during the contextual fear memory was significantly lowered in stressed females and males (figure 3I, supplemental figure 1A-B). The auditory fear retrieval did not show a different freezing response on mean tones in females (figure 3K, supplemental figure 1C), but did show a lowered freezing response in LBN males compared to NS (figure 3K, supplemental figure 1D). In addition, no differences were observed in the fear retrieval for both females and males (figure 3L, supplemental figure 1E-F).

The DeepOF unsupervised clustering analysis of the fear acquisition data yielded 9 distinct clusters (figure 4A, B). No cluster population differences were observed by the stress background for both female and male ITI 1-4 data (supplemental figure 2 A, B). However, clusters were significantly altered by the stress background in females during tones 2 to 5 (figure 4A), which were not observed in males (figure 4B). In particular, “cluster 0” was significantly increased in LBN females, and “cluster 6” significantly decreased in LBN females compared to NS (figure 4A). A multi-class supervised learning model was trained to map from motion summary statistics to the obtained cluster labels, and performance was measured in terms of the balanced accuracy per cluster (figure 4C). The confusion matrix showed low probabilities for all cluster crossovers and the classifier performance was substantially greater than random for all clusters, indicating that all clusters were substantially distinguishable by the model (figure 4C). The cluster detection analysis yields a set of feature explainers per cluster that can be used to interpret the clusters using SHAP values in global (figure 4D) and cluster-specific ways (figure 4E, F). Importantly, the interpretation of the clusters was done using the feature importance of the SHAP analysis, together with the visual interpretation of the video fragments per cluster (see supplemental materials for video output per cluster). The global feature importance across all clusters revealed that the distance towards several spine labels (a stretch or a shortening of the back), the overall speed (an increased or reduced speed), huddle (an increased or decreased amount of the behavior in which the animal stops moving around and

bends the back), and the surface area of the head (an increased surface area is related to the head being forward, whereas a decreased surface area is related to the head being downward) were particularly important for global cluster inclusion (figure 4D). More specifically, the feature importance analysis for “cluster 0” revealed that an increased speed, a decrease in huddle behavior, and an increased spine stretch were important features for cluster inclusion (figure 4E). The visual inspection of “cluster 0” indicated a behavior related to the exploration of the environment, in which the animal was moving around (see supplemental materials). In contrast, “cluster 6” feature importance analysis revealed that a decrease in spine stretch, an increase in huddle behavior, and a reduction of the surface head area were important features for cluster inclusion (figure 4F). The visual inspection of “cluster 6” indicated a behavior related to freezing, in which the animal was often immobile and close to the outside of the environment (see supplemental materials).

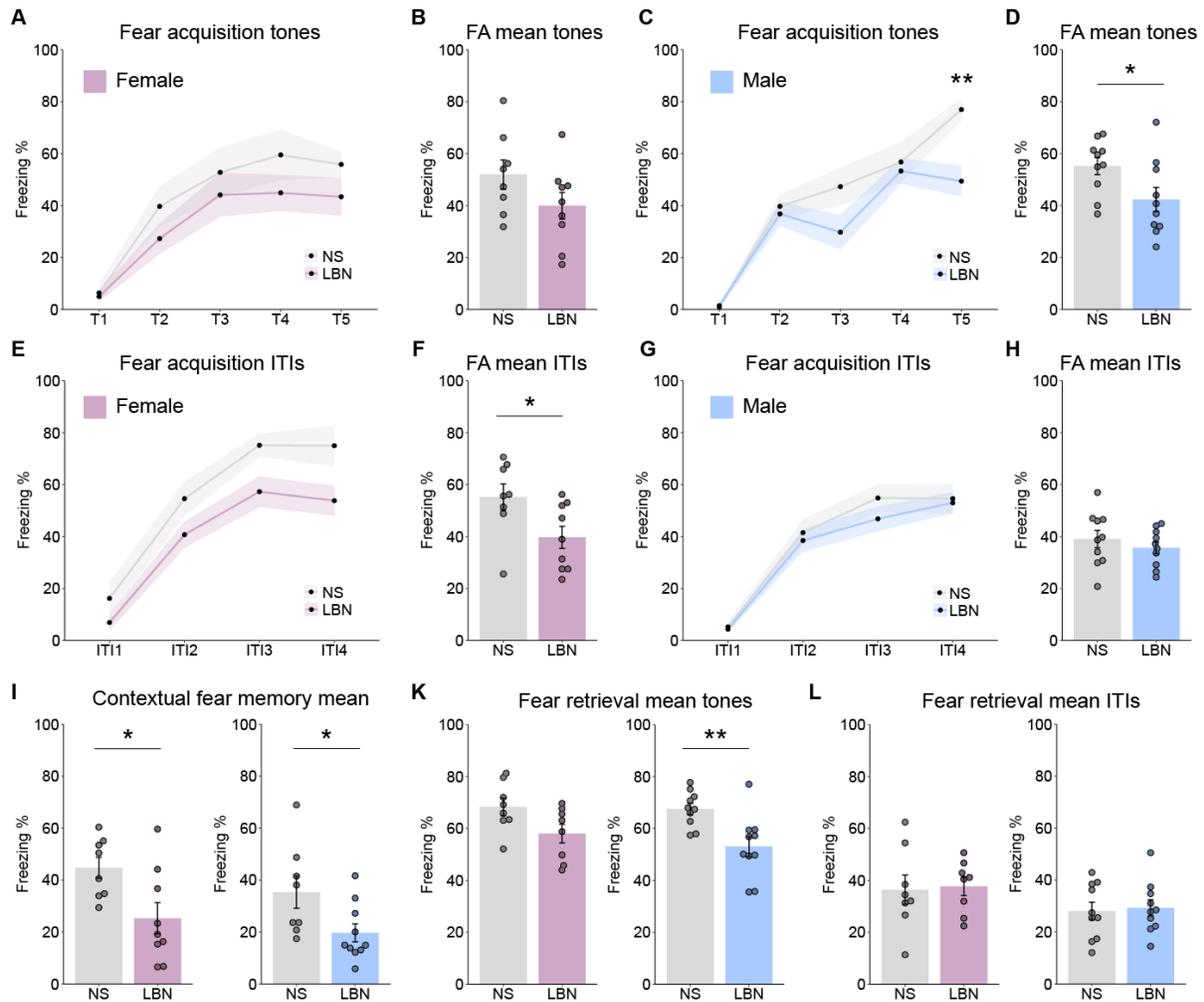


Figure 3. Fear conditioning data on freezing behavior. A) Freezing behavior on individual tones during fear acquisition (FA) was

not altered between nonstressed and stressed females, two-way ANOVA ($p > 0.45$) **B**) The average freezing during the FA tones 2-5 was not altered between nonstressed and stressed females ($T(15) = 1.58$, $p = 0.14$). **C**) Freezing behavior in the FA for individual tones 1-4 was not altered between nonstressed and stressed males ($p > 0.22$), but for tone 5 it was significantly lowered in stressed males ($F(1,18) = 14.3$, $p = 0.005$), with a significant main effect for the two-way ANOVA on stress ($F(1,90) = 9.39$, $p = 0.003$), tones ($F(4,90) = 41.30$, $p < 0.0001$), and stress*tones ($F(4,90) = 2.6$, $p = 0.041$). **D**) The average freezing during FA tones 2-5 was significantly lowered in stressed males compared to nonstressed ($T(18) = 2.25$, $p = 0.037$). **E**) No significant main effect could be observed between nonstressed and stressed females for the individual ITIs in the FA. **F**) The average freezing during FA ITIs 1-4 was significantly lowered in stressed females compared to nonstressed ($T(15) = 2.37$, $p = 0.03$). **G**) No significant main effect could be observed between nonstressed and stressed males for the individual ITIs in the FA. **H**) The average freezing during FA ITIs 1-4 was not altered between nonstressed and stressed males ($T(18) = 0.84$, $p = 0.41$). **I**) The mean freezing in the contextual fear memory task was significantly lowered in stressed females compared to nonstressed ($T(15) = 2.61$, $p = 0.020$), and in males ($T(16) = 2.32$, $p = 0.034$). **K**) The mean freezing in the fear retrieval task showed no significant difference in females ($T(14) = 2.09$, $p = 0.055$). However, in males, a significant reduction in freezing was observed in stressed compared to nonstressed ($T(18) = 3.29$, $p = 0.004$). **L**) No significant differences were observed in the mean freezing during the ITIs between stressed and nonstressed animals for females and males.

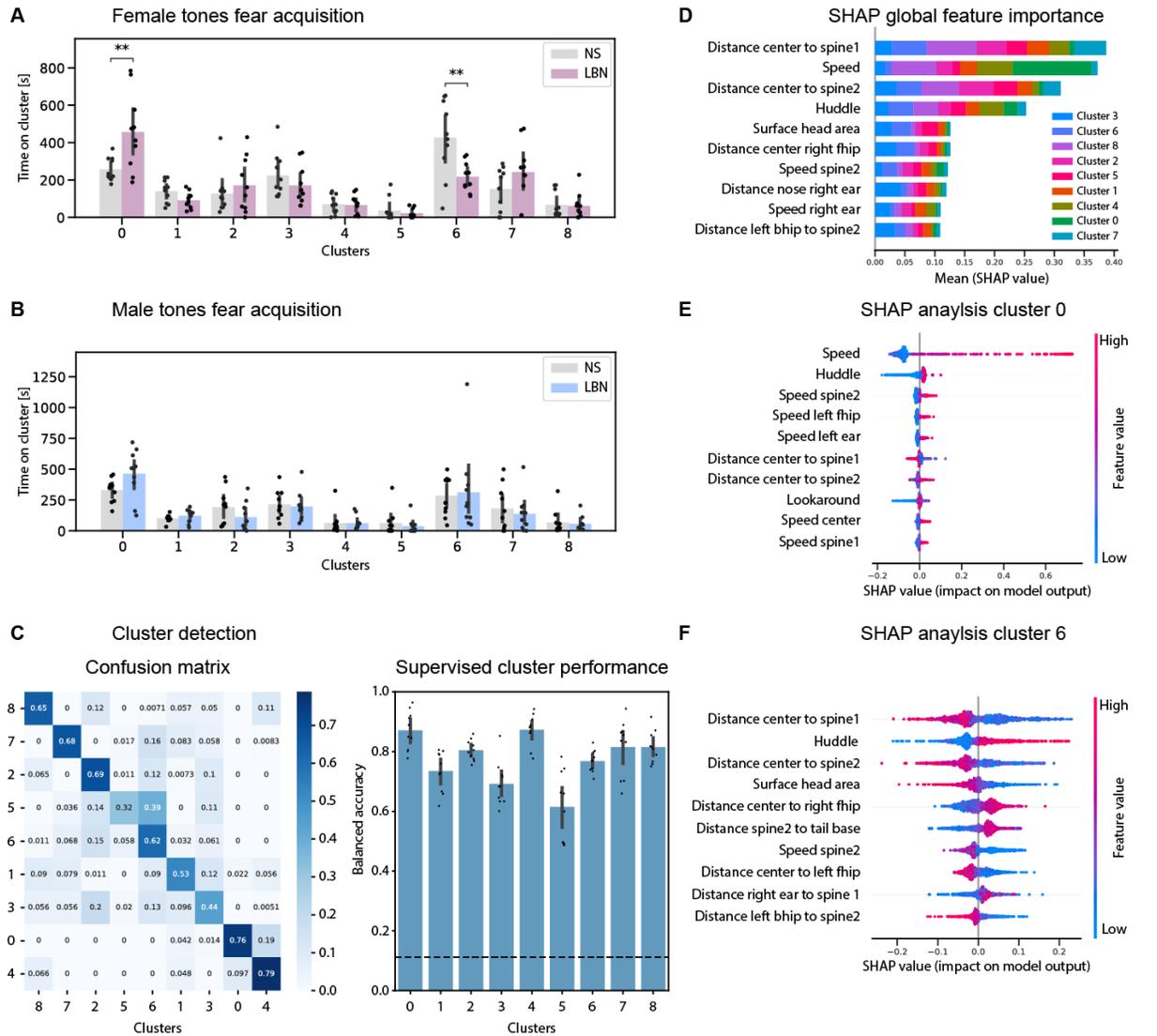


Figure 4. Unsupervised analysis of fear acquisition data on tones 2-5. A) Cluster enrichment for the female fear acquisition data using tones 2-5. Bar graphs represent mean \pm standard deviation of the time proportion spent on each cluster. Statistics are performed using an independent samples t-test corrected for multiple testing using Benjamini-Hochberg's method across clusters. Significant differences were observed in clusters 0: $T=-3.55$, $p=2.28 \times 10^{-3}$, and 6: $T=4.24$, $p=4.97 \times 10^{-4}$, but none of the other clusters ($p>0.05$). **B)** Cluster enrichment for the male fear acquisition data using tones 2-5. No significant differences were observed using the independent samples t-test corrected for multiple testing using Benjamini-Hochberg's method across clusters ($p>0.05$). Bar graphs represent mean \pm standard deviation of the time proportion spent on each cluster. **C)** On the left, the confusion matrix is obtained from the trained gradient boosting machine classifying between clusters. Aggregated performance over the validation folds of 10-fold cross-validation is shown. On the right, is the validation performance per cluster across a 10-fold cross-validation loop. Balanced accuracy was used to correct for cluster assignment imbalance. The dashed line marks are the expected performance due to chance, considering all outputs. **D)** The global SHAP feature importance between the different clusters. Features in the y-axis are sorted on the global absolute SHAP values across all clusters. The classes in the bar graphs are sorted by highest to lowest clusters importance within every feature. **E-F)** Bee swarm plots for the two differentially expressed

clusters within the female fear acquisition data between NS and LBN mice, clusters 0 and 6. The plots show the 10 most important features for each classifier, in terms of the mean absolute value of the SHAP values.

Discussion

Exposure to ELS increases the vulnerability toward stress-related disorders, such as PTSD. The prevalence of PTSD is strongly influenced by sex. Animal models on ELS exposure have found an increasing body of evidence that LBN exposure affects rodents in a sex-specific manner^{40–44}. However, the sex-dependent effects of HPA-axis signaling and fear memory formation on ELS have yet to be uncovered. The development of open-source markerless pose estimation tools⁵¹ and subsequently unsupervised behavioral analysis tools has allowed for in-depth behavioral analysis that can explore previously unknown behavioral patterns^{50,52–54}. This is crucial to increase the understanding of the behavioral outcome related to stress-induced fear memory formation. The current study explores HPA axis mechanisms in the body and brain and the behavioral output related to LBN-induced fear memory formation in a sex- and time-dependent manner.

LBN disrupts different facets of the HPA axis in a time- and sex-dependent manner

The LBN model has been extensively utilized to investigate the effects of chronic ELS exposure on both physiological and behavioral outcomes. A common hallmark of LBN exposure is the reduction of body weight at P09⁴⁷, which was confirmed in the current study, highlighting that the LBN model has stress-dependent effects in both sexes^{55,56}. In addition, the long-term effect of LBN on body weight between sexes is more variable and seems to be dependent on the age of testing and potential additional challenges throughout adulthood. The current study showed that at 3 months of age, females show a more persistent LBN-related body weight reduction phenotype compared to males. Other studies have shown a similar effect at 2 months of age⁵⁵, but Arp et al. did not find this sex-dependent difference in 4 months old animals⁵⁷, indicating that both males and females eventually recover, but females show a longer recovery period. An opposite effect was observed in another study at 8 months of age, where males showed a stronger body weight reduction compared to females⁵⁶, this might be explained by the different adult stress events, (e.g. glucose- and insulin tolerance tests) and indicates different vulnerability toward such adult stressors after LBN exposure between sexes.

ELS exposure has been linked to dysregulation of the HPA axis, which can lead to an increased vulnerability state of stress-related psychopathology^{21,22}. A well-established phenomenon in rats is the elevated levels of morning CORT baseline in females compared to males^{58–60}. The current study replicates this phenomenon in adult mice regardless of stress exposure, as also recently reported by Brix et al.⁵⁶. Interestingly, it was further observed that this sex-dependent difference is already apparent at the early age of P09, highlighting that the sex-specific differences are apparent already at the end of ELS exposure.

An earlier study on LBN exposure in males showed a significant LBN-induced increase at P09 for baseline CORT in mice ⁴⁷, but the current study found an increase in CORT only in LBN females. This might be explained due to the low baseline levels of CORT in males, which therefore might show a higher variance, as the absolute CORT values between conditions are smaller. Nonetheless, the high increase of baseline CORT in LBN-exposed females is a good proxy for stress exposure, and at P09 is indicated to be higher in females compared to males. Moreover, we observed an opposite effect for adrenal weight at adult age, which was significantly increased for LBN-exposed males, but not females. This is in line with earlier research, that showed a similar effect in males at 1 month of age ⁵⁷, but at later stages in adulthood, namely 4 months and 8 months, the adrenals in males were back to the same size as the nonstressed condition ^{47,56}. This indicates that the adrenal size is influenced in a time-dependent manner, in which males are taking longer to recover their adrenal size to baseline after LBN exposure.

Another facet of the HPA-axis reactivity was investigated via gene expression changes of *Fkbp5* in the brain. Previous research has identified a particularly high expression of the *Fkbp5* gene in the BLA and HIP under baseline ³⁸. This study replicates the high *Fkbp5* gene expression pattern in the BLA and HIP at P56, in which the DG shows the highest expression from the HIP subregions during baseline, as similarly observed by Scharf et al. ³⁸. Furthermore, we show that *Fkbp5* gene expression under baseline can already be observed in the BLA and HIP at P09, during which the CA2-3 subregion shows a higher *Fkbp5* expression than the DG, indicating an age-dependent expression pattern of *Fkbp5* in the subregions of the dorsal HIP. Moreover, *Fkbp5* gene expression has been shown to increase in a stressor-dependent manner in the BLA and HIP ^{38,61,62}. However, the immediate and long-term *Fkbp5* gene expression changes in response to LBN have remained elusive. We show that *Fkbp5* gene expression was not changed by LBN exposure directly after the stress at P09 in both sexes but was upregulated specifically in the CA1 region of the dorsal HIP of adult LBN-exposed males, but not females. Marrero et al. 2019 ³⁹ showed that the overexpression of human FKBP5 in the forebrain induces specific downstream molecular changes in the dorsal HIP in adult ELS animals using the maternal separation paradigm. This coincides with the current finding that specifically the dorsal HIP shows upregulated *Fkbp5* expression and points to an altered molecular pathway mechanism after LBN exposure. In conclusion, we highlight a differential impact of LBN exposure across sexes. The immediate effects of LBN exposure at P09 are more pronounced in females, while interestingly, the prolonged effects of a dysregulated HPA-axis in adult age are affected exclusively in males.

Fear acquisition is differentially affected by LBN across sexes

The formation of fear memory is a crucial aspect of understanding the underlying mechanism of PTSD. The specific alterations of anxiety and fear behavior can be investigated using animal models of fear conditioning. In line with previous research, we show that exposure to LBN reduces the fear response by lowering freezing behavior during both contextual- and auditory fear retrieval in males⁴⁶. In addition, we show that LBN exposure in females shows a similar reduction in freezing behavior during the contextual fear retrieval, but not in the auditory fear retrieval. Previous research has shown that LBN is linked to reduced synaptic plasticity markers within the dorsal HIP, which could explain the LBN-induced differences in fear retrieval by altering fear memory formation⁴⁶. However, the current study highlights an alternative explanation by specifically exploring the fear behavior during the acquisition of fear conditioning. Specifically, it was observed that the reduction in freezing behavior can already be found during the acquisition of the fear memory, in which LBN-exposed males show reduced freezing during the acquisition tones, and the females during the acquisition ITIs. A similar effect has been observed in other studies for male data, in which it was shown that the freezing directly after the fear acquisition is already lowered in LBN-exposed animals^{46,57,63}. Therefore, the difference in freezing during the retrieval phase is not only explained by differences in fear memory formation but also by an altered response at fear acquisition.

In addition, several studies have highlighted the relevance of distinguishing between different types of fear behaviors^{64–66}. The analysis of the behavioral data using an unsupervised analysis provides a promising way to explore novel behavioral patterns related to fear acquisition without prior behavioral categorization. This allows the exploration of the behavioral repertoire in a hypothesis-generating way, which can lead to the identification of novel behaviors within the specific methodological context⁶⁷. To further understand the sex-dependent fear behavioral differences during fear acquisition, the DeepOF open source python package was deployed to perform an unsupervised analysis pipeline, which maps the representations of different fear behavior-related syllables across the different stages (tones and ITIs), conditions (NS vs LBN) and sex (female vs male) without any prior label information. Different fear-related behaviors were observed that were particularly altered in female, but not male mice. Interestingly, it was observed that specific behavioral clusters (e.g., cluster 0) are elevated in LBN-exposed females, which indicated a behavior related to the exploration of the environment. Conversely, other behavioral clusters (e.g., cluster 6) are reduced in LBN-exposed females, which were related to freezing behavior. Intriguingly, the observed behavior in “cluster 0” coincides with a previously identified active fear behavioral response,

called “darting”, in which rapid locomotive movements are detected in primarily female rats ⁶⁶. The behavioral syllables from “cluster 0” can also be allocated to an active fear behavioral response, but under nonstressed conditions are expressed in both females as well as males. However, the increased amount of the expression of “cluster 0” after LBN exposure is exclusively observed in females, which does indicate a sex-dependent effect on the active fear response.

Conclusion

Taken together, the current study shows a sex-specific effect of LBN exposure on dysregulation of the HPA-axis, in which the adrenal weight, baseline CORT levels, and *Fkbp5* expression in several stress-related brain regions, including the BLA and subregions of the dorsal HIP, are altered in a time-dependent manner. In addition, we show that specific aspects of fear-related behavior, including the passive fear behavioral response via freezing behavior, but also an active fear response, as identified using an unsupervised analysis, are altered by LBN exposure in a sex-specific manner. The additional fear-related behavior that is expressed in “cluster 0” is contributing to a better understanding of the sex-dependent effects of fear memory acquisition and might influence the expression of the freezing behavior during contextual as well as auditory fear retrieval. The DeepOF unsupervised analysis provides an additional layer to explore the fear-related behaviors without prior assumptions and therefore allows for hypothesis-generating behavioral analysis, which ultimately can lead to a better understanding of the stress-induced behavioral phenotype.

Acknowledgements

The authors thank Lisa Rudolph, Daniela Harbich, and Bianca Schmid for their technical assistance, and the DeepLabCut development team for creating and maintaining the DeepLabCut software.

Author contributions

JB and MVS conceived the study. JB performed the experiments, LvD, LMB, SN, HY, SM, VK and MS assisted with the experiments. JB analyzed the data and was assisted by LM and BMM.

Funding

This study is supported by the “Kids2Health” grant of the Federal Ministry of Education and Research [01GL1743C].

References

1. Pervanidou, P. & Chrousos, G. P. Early-Life Stress: From Neuroendocrine Mechanisms to Stress-Related Disorders. *Horm Res Paediatr* **89**, 372–379 (2018).
2. Agorastos, A., Pervanidou, P., Chrousos, G. P. & Kolaitis, G. Early life stress and trauma: developmental neuroendocrine aspects of prolonged stress system dysregulation. *Hormones* **17**, 507–520 (2018).
3. Ridout, K. K. *et al.* Molecular markers of neuroendocrine function and mitochondrial biogenesis associated with early life stress. *Psychoneuroendocrinology* **116**, 104632 (2020).
4. Saleh, A. *et al.* Effects of early life stress on depression, cognitive performance and brain morphology. *Psychol Med* **47**, 171–181 (2017).
5. Teicher, M. H. & Samson, J. A. Annual Research Review: Enduring neurobiological effects of childhood abuse and neglect. *Journal of Child Psychology and Psychiatry* **57**, 241–266 (2016).
6. Heim, C. & Nemeroff, C. B. The role of childhood trauma in the neurobiology of mood and anxiety disorders: preclinical and clinical studies. *Biol Psychiatry* **49**, 1023–1039 (2001).
7. Nemeroff, C. B. Paradise Lost: The Neurobiological and Clinical Consequences of Child Abuse and Neglect. *Neuron* **89**, 892–909 (2016).
8. Pervanidou, P., Makris, G., Chrousos, G. & Agorastos, A. Early Life Stress and Pediatric Posttraumatic Stress Disorder. *Brain Sci* **10**, 169 (2020).
9. Kohli, M. A. *et al.* The Neuronal Transporter Gene SLC6A15 Confers Risk to Major Depression. *Neuron* **70**, 252–265 (2011).
10. Howard, D. M. *et al.* Genome-wide haplotype-based association analysis of major depressive disorder in Generation Scotland and UK Biobank. *Transl Psychiatry* **7**, 1263 (2017).
11. Gałecka, E. *et al.* Single nucleotide polymorphisms of NR3C1 gene and recurrent depressive disorder in population of Poland. *Mol Biol Rep* **40**, 1693–1699 (2013).
12. Klok, M. D. *et al.* Decreased expression of mineralocorticoid receptor mRNA and its splice variants in postmortem brain regions of patients with major depressive disorder. *J Psychiatr Res* **45**, 871–878 (2011).
13. Caspi, A. *et al.* Influence of Life Stress on Depression: Moderation by a Polymorphism in the 5-HTT Gene. *Science (1979)* **301**, 386–389 (2003).
14. van West, D. *et al.* A major SNP haplotype of the arginine vasopressin 1B receptor protects against recurrent major depression. *Mol Psychiatry* **9**, 287–292 (2004).
15. Binder, E. B. *et al.* Polymorphisms in FKBP5 are associated with increased recurrence of depressive episodes and rapid response to antidepressant treatment. *Nat Genet* **36**, 1319–1325 (2004).

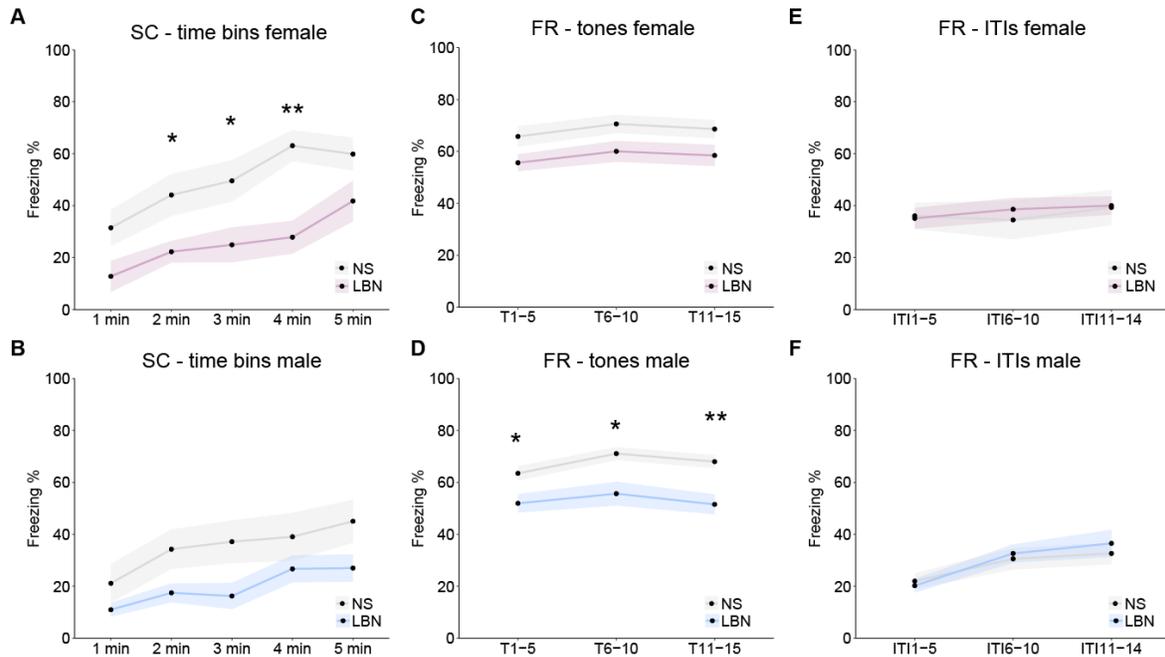
16. Binder, E. B. *et al.* Association of FKBP5 polymorphisms and childhood abuse with risk of posttraumatic stress disorder symptoms in adults. *Journal of American Medical Association* **299**, 1291–1305 (2008).
17. Klengel, T. *et al.* Allele-specific FKBP5 DNA demethylation mediates gene – childhood trauma interactions. *Nat Neurosci* **16**, (2013).
18. Kang, J. I., Kim, T. Y., Choi, J. H., So, H. S. & Kim, S. J. Allele-specific DNA methylation level of FKBP5 is associated with post-traumatic stress disorder. *Psychoneuroendocrinology* **103**, 1–7 (2019).
19. Xie, P. *et al.* Interaction of FKBP5 with Childhood Adversity on Risk for Post-Traumatic Stress Disorder. *Neuropsychopharmacology* **35**, 1684–1692 (2010).
20. Baker, J., Ozsan, I., Rodriguez Ospina, S., Gulick, D. & Blair, L. Hsp90 Heterocomplexes Regulate Steroid Hormone Receptors: From Stress Response to Psychiatric Disease. *Int J Mol Sci* **20**, 79 (2018).
21. de Kloet, E. R., Joëls, M. & Holsboer, F. Stress and the brain: from adaptation to disease. *Nat Rev Neurosci* **6**, 463–475 (2005).
22. Juruena, M. F., Eror, F., Cleare, A. J. & Young, A. H. The Role of Early Life Stress in HPA Axis and Anxiety. in 141–153 (2020). doi:10.1007/978-981-32-9705-0_9.
23. Kilpatrick, D. G. *et al.* National Estimates of Exposure to Traumatic Events and PTSD Prevalence Using *DSM-IV* and *DSM-5* Criteria. *J Trauma Stress* **26**, 537–547 (2013).
24. Olf, M., Langeland, W., Draijer, N. & Gersons, B. P. R. Gender differences in posttraumatic stress disorder. *Psychol Bull* **133**, 183–204 (2007).
25. Bangasser, D. A. & Valentino, R. J. Sex differences in stress-related psychiatric disorders: Neurobiological perspectives. *Front Neuroendocrinol* **35**, 303–319 (2014).
26. Teicher, M. H., Anderson, C. M. & Polcari, A. Childhood maltreatment is associated with reduced volume in the hippocampal subfields CA3, dentate gyrus, and subiculum. *Proceedings of the National Academy of Sciences* **109**, (2012).
27. Frodl, T., Reinhold, E., Koutsouleris, N., Reiser, M. & Meisenzahl, E. M. Interaction of childhood stress with hippocampus and prefrontal cortex volume reduction in major depression. *J Psychiatr Res* **44**, 799–807 (2010).
28. Humphreys, K. L. *et al.* Evidence for a sensitive period in the effects of early life stress on hippocampal volume. *Dev Sci* **22**, (2019).
29. Burghy, C. A. *et al.* Developmental pathways to amygdala-prefrontal function and internalizing symptoms in adolescence. *Nat Neurosci* **15**, 1736–1741 (2012).
30. Vythilingam, M. *et al.* Childhood Trauma Associated With Smaller Hippocampal Volume in Women With Major Depression. *American Journal of Psychiatry* **159**, 2072–2080 (2002).

31. Pleasure, S. J., Collins, A. E. & Lowenstein, D. H. Unique Expression Patterns of Cell Fate Molecules Delineate Sequential Stages of Dentate Gyrus Development. *The Journal of Neuroscience* **20**, 6095–6105 (2000).
32. Sorrells, S. F. *et al.* Immature excitatory neurons develop during adolescence in the human amygdala. *Nat Commun* **10**, 2748 (2019).
33. Avino, T. A. *et al.* Neuron numbers increase in the human amygdala from birth to adulthood, but not in autism. *Proceedings of the National Academy of Sciences* **115**, 3710–3715 (2018).
34. Galea, L. A. M., Leuner, B. & Slattery, D. A. Hippocampal Plasticity during the Peripartum Period: Influence of Sex Steroids, Stress and Ageing. *J Neuroendocrinol* **26**, 641–648 (2014).
35. Lupien, S. J., McEwen, B. S., Gunnar, M. R. & Heim, C. Effects of stress throughout the lifespan on the brain, behaviour and cognition. *Nat Rev Neurosci* **10**, 434–445 (2009).
36. Geuze, E. *et al.* Glucocorticoid receptor number predicts increase in amygdala activity after severe stress. *Psychoneuroendocrinology* **37**, 1837–1844 (2012).
37. Han, F., Ding, J. & Shi, Y. Expression of amygdala mineralocorticoid receptor and glucocorticoid receptor in the single-prolonged stress rats. *BMC Neurosci* **15**, 77 (2014).
38. Scharf, S. H., Liebl, C., Binder, E. B., Schmidt, M. V. & Müller, M. B. Expression and regulation of the Fkbp5 gene in the adult mouse brain. *PLoS One* **6**, 1–10 (2011).
39. Criado-Marrero, M. *et al.* Early Life Stress and High FKBP5 Interact to Increase Anxiety-Like Symptoms through Altered AKT Signaling in the Dorsal Hippocampus. *Int J Mol Sci* **20**, 2738 (2019).
40. Goodwill, H. L. *et al.* Early life stress leads to sex differences in development of depressive-like outcomes in a mouse model. *Neuropsychopharmacology* **44**, 711–720 (2019).
41. Laham, B. J. *et al.* The estrous cycle modulates early-life adversity effects on mouse avoidance behavior through progesterone signaling. *Nat Commun* **13**, 7537 (2022).
42. White, J. D. *et al.* Early life stress causes sex-specific changes in adult fronto-limbic connectivity that differentially drive learning. *Elife* **9**, (2020).
43. Bachiller, S. *et al.* Early-life stress elicits peripheral and brain immune activation differently in wild type and 5xFAD mice in a sex-specific manner. *J Neuroinflammation* **19**, 151 (2022).
44. lo Iacono, L. *et al.* Early life adversity affecting the attachment bond alters ventral tegmental area transcriptomic patterning and behavior almost exclusively in female mice. *Neurobiol Stress* **15**, 100406 (2021).
45. Bonapersona, V. *et al.* The mouse brain after foot shock in four dimensions: Temporal dynamics at a single-cell resolution. *Proceedings of the National Academy of Sciences* **119**, (2022).
46. Lesuis, S. L., Lucassen, P. J. & Krugers, H. J. Early life stress impairs fear memory and synaptic plasticity; a potential role for GluN2B. *Neuropharmacology* **149**, 195–203 (2019).

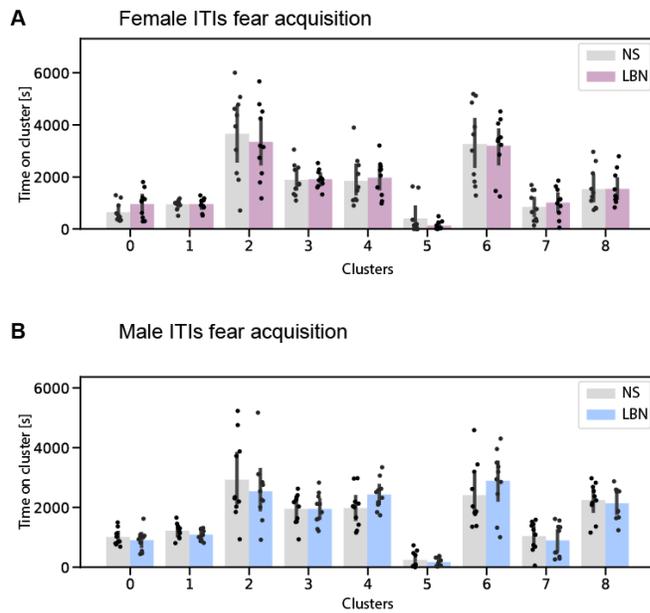
47. Rice, C. J., Sandman, C. A., Lenjavi, M. R. & Baram, T. Z. A novel mouse model for acute and long-lasting consequences of early life stress. *Endocrinology* **149**, 4892–4900 (2008).
48. Hartmann, J. *et al.* Forebrain glutamatergic, but not GABAergic, neurons mediate anxiogenic effects of the glucocorticoid receptor. *Mol Psychiatry* **22**, 466–475 (2017).
49. Lauer, J. *et al.* Multi-animal pose estimation, identification and tracking with DeepLabCut. *Nat Methods* **19**, 496–504 (2022).
50. Bordes, J. *et al.* Automatically annotated motion tracking identifies a distinct social behavioral profile following chronic social defeat stress. *bioRxiv* 2022.06.23.497350 (2022) doi:10.1101/2022.06.23.497350.
51. Mathis, A. *et al.* DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat Neurosci* **21**, 1281–1289 (2018).
52. Yttri, E., Hsu, A., Schweihoff, J. & Schwarz, M. A-SOiD, an active learning platform for expert-guided, data efficient discovery of behavior. (2023) doi:10.21203/rs.3.rs-2239160/v1.
53. Weinreb, C. *et al.* Keypoint-MoSeq: parsing behavior by linking point tracking to pose dynamics. *bioRxiv* 2023.03.16.532307 (2023) doi:10.1101/2023.03.16.532307.
54. Luxem, K. *et al.* Identifying behavioral structure from deep variational embeddings of animal motion. *Commun Biol* **5**, 1267 (2022).
55. Bonapersona, V. *et al.* Sex-Dependent Modulation of Acute Stress Reactivity After Early Life Stress in Mice: Relevance of Mineralocorticoid Receptor Expression. *Front Behav Neurosci* **13**, (2019).
56. Brix, L. M. *et al.* Metabolic effects of early life stress and pre-pregnancy obesity are long-lasting and sex-specific in mice. *European Journal of Neuroscience* (2023) doi:10.1111/ejn.16047.
57. Arp, J. M. *et al.* Blocking glucocorticoid receptors at adolescent age prevents enhanced freezing between repeated cue-exposures after conditioned fear in adult mice raised under chronic early life stress. *Neurobiol Learn Mem* **133**, 30–38 (2016).
58. Critchlow, V., Liebelt, R. A., Bar-Sela, M., Mountcastle, W. & Lipscomb, H. S. Sex difference in resting pituitary-adrenal function in the rat. *American Journal of Physiology-Legacy Content* **205**, 807–815 (1963).
59. Seale, J. V. *et al.* Gonadectomy Reverses The Sexually Diergic Patterns Of Circadian and Stress-Induced Hypothalamic-Pituitary-Adrenal Axis Activity In Male and Female Rats. *J Neuroendocrinol* **16**, 516–524 (2004).
60. Spiga, F., Walker, J. J., Terry, J. R. & Lightman, S. L. HPA Axis-Rhythms. in *Comprehensive Physiology* 1273–1298 (Wiley, 2014). doi:10.1002/cphy.c140003.
61. Guidotti, G. *et al.* Glucocorticoid receptor and fkbp5 expression is altered following exposure to chronic stress: Modulation by antidepressant treatment. *Neuropsychopharmacology* **38**, 616–627 (2013).

62. Santarelli, S. *et al.* An adverse early life environment can enhance stress resilience in adulthood. *Psychoneuroendocrinology* **78**, 213–221 (2017).
63. Kanatsou, S. *et al.* Overexpression of Mineralocorticoid Receptors in the Mouse Forebrain Partly Alleviates the Effects of Chronic Early Life Stress on Spatial Memory, Neurogenesis and Synaptic Function in the Dentate Gyrus. *Front Cell Neurosci* **11**, 1–13 (2017).
64. Mitchell, J. R. *et al.* Darting across space and time: parametric modulators of sex-biased conditioned fear responses. *Learning & Memory* **29**, 171–180 (2022).
65. Fadok, J. P. *et al.* A competitive inhibitory circuit for selection of active and passive fear responses. *Nature* **542**, 96–100 (2017).
66. Gruene, T. M., Flick, K., Stefano, A., Shea, S. D. & Shansky, R. M. Sexually divergent expression of active and passive conditioned fear responses in rats. *Elife* **4**, (2015).
67. Bordes, J., Miranda, L., Müller-Myhsok, B. & Schmidt, M. V. Advancing social behavioral neuroscience by integrating ethology and comparative psychology methods through machine learning. *Neurosci Biobehav Rev* **151**, 105243 (2023).

Supplemental materials



Supplemental figure 1. Timebin data for contextual and retrieval fear memory. A) The freezing behavior during 1min bins of the contextual fear memory in females. A significant main effect was observed using the two-way ANOVA on stress ($F(1,85)=31.61$, $p<0.0001$), and tones ($F(4,85)=5.54$, $p=0.0005$), but not on stress*tones ($p=0.70$). Post-hoc analysis using BH revealed a significant reduction of freezing in LBN females compared to NS for 2min ($p=0.047$), 3min ($p=0.047$), and 4min ($p=0.004$), but not 1min ($p=0.069$), and 5min ($p=0.094$). **B)** No significant main effect was observed using the Kruskal Wallis test ($p>0.15$). **C)** The freezing behavior during the auditory fear retrieval tones binned per 5 tones in females. A significant main effect was observed using the two-way ANOVA on stress ($F(1,42)=11.59$, $p=0.001$), but not for tones, or stress*tones ($p>0.46$). Post-hoc analysis using BH revealed no further significance between stress conditions ($p=0.077$). **D)** The freezing behavior during the auditory fear retrieval tones binned per 5 tones in males. A significant main effect was observed using the two-way ANOVA on stress ($F(1,54)=27.96$, $p<0.0001$), but not for tones, or stress*tones ($p>0.24$). Post-hoc analysis using BH revealed a significantly lowered freezing response in LBN males compared to NS at T1-5 ($F(1,18)=6.695$, $p=0.019$), T6-10 ($F(1,18)=8.84$, $p=0.012$), and T11-15 ($F(1,18)=13.12$, $p=0.006$). **E)** No significant main effect was observed using the two-way ANOVA for the fear retrieval ITIs in females ($p>0.73$). **F)** No significant differences were observed between LBN and NS males between the different ITIs in the fear retrieval task; the two-way ANOVA did reveal a significant main effect for ITIs ($F(2,54)=6.80$, $p=0.002$), but not for stress ($F(1,54)=0.20$, $p=0.66$), or stress*ITIs ($F(2,54)=0.28$, $p=0.76$). Further post-hoc analysis using BH revealed no significant differences ($p>0.57$).



Supplemental figure 2. Unsupervised clusters during the ITIs. A) Cluster enrichment for the female fear acquisition data using all four ITIs. No significant differences were observed using the independent samples t-test corrected for multiple testing using Benjamini-Hochberg’s method across clusters ($p > 0.05$). **B)** Cluster enrichment for the male fear acquisition data using all four ITIs. No significant differences were observed using the independent samples t-test corrected for multiple testing using Benjamini-Hochberg’s method across clusters ($p > 0.05$). Bar graphs represent mean \pm standard deviation of the time proportion spent on each cluster.

Chapter 6

The Fkbp5 knock-out in the noradrenergic locus coeruleus system alters social behavior in male mice

Bordes, J., Bajaj, T., Chang, S., Miranda, L., Pöhlmann, M., Meccad, Y., Anderzhanova, E., Brix, L.M., van Doeselaar, L., Narayan, S., Yang, H., Mitra, S., Kovarova, V., Springer, M., de Angelis, M., Müller-Myhsok, B., Deussing, J., Gassen, N., Schmidt, M.V.

Manuscript in preparation

The *Fkbp5* knock-out in the noradrenergic locus coeruleus system alters social behavior in male mice

Joeri Bordes¹, Thomas Bajaj², Simon Chang³, Lucas Miranda^{4,5}, Max Pöhlmann¹, Yara Meccad², Elmira Anderzhanova², Lea Maria Brix^{1,5}, Lotte van Doeselaar^{1,5}, Sowmya Narayan^{1,5}, Huanqing Yang¹, Shiladitya Mitra¹, Veronika Kovarova^{1,5}, Margherita Springer¹, Meri de Angelis⁶, Bertram Müller-Myhsok⁴, Jan Deussing³, Nils Gassen², Mathias V. Schmidt¹

Affiliations:

¹Research Group Neurobiology of Stress Resilience, Max Planck Institute of Psychiatry, 80804 Munich, Germany

²Neurohomeostasis Research Group, Department of Psychiatry and Psychotherapy, Bonn Clinical Center, University of Bonn, 53127

Bonn, Germany.

³ Research Group Neurogenetics, Max Planck Institute of Psychiatry, Munich, Germany

⁴Research Group Statistical Genetics, Max Planck Institute of Psychiatry, 80804 Munich, Germany

⁵International Max Planck Research School for Translational Psychiatry (IMPRS-TP), 80804 Munich, Germany

⁶Helmholtz Center Munich Germany Research Center for Environmental Health, Molecular Exposomics, Neuherberg, Germany.

Abstract

Social interactions play a crucial role in our daily lives, and their dysregulation is often observed in psychiatric disorders such as major depressive disorder (MDD). The vulnerability to MDD is influenced by a combination of genetic, epigenetic, and environmental factors. A particular gene of interest in MDD research is FKBP5, which encodes for the co-chaperone FK506 binding protein 51 (FKBP51). The interaction between severe stress exposure and genetic risk variants of FKBP5 has been associated with an increased risk for MDD vulnerability. The region-specific role of FKBP51 and its impact on social functioning remains to be fully elucidated. The current study observed a region-specific role of *Fkbp5* regulation within the locus coeruleus (LC) in response to different types of stressors. The results showed that only acute social stress led to a significant upregulation of *Fkbp5* in the LC, highlighting the importance of the social nature of stressors in *Fkbp5* regulation within the LC. Furthermore, the conditional knockout of *Fkbp5* within the noradrenergic system (*Fkbp5*^{Nat}) impacts the social behavioral profile towards a novel social conspecific in a sex-dependent manner, as identified only in male mice using a deep phenotyping strategy via a supervised behavioral analysis. This sex-dependent effect suggests that LC *Fkbp5* plays a role in modulating male social behavior. Furthermore, the study reveals that *Fkbp5*^{Nat} mice show long-lasting changes in the molecular pathway related to the noradrenergic synapse under baseline and altered noradrenergic signaling when exposed to a novel social stimulus, specifically in the basolateral amygdala region. Overall, this study contributes to our understanding of the complex interplay between genetic and environmental factors in psychiatric disorders, shedding light on the role of *Fkbp5* and the noradrenergic system in the regulation of social behavior and its implications for psychiatric risk exposure.

Introduction

Social interactions play a crucial role in our daily lives and have a crucial influence on the determination of many societal events, such as making friendships, finding love, and finding a successful career path. However, the expression of social behavior and the response to social challenges has a high inter-individual variability, depending on a heterogeneous expression of emotional and cognitive-related behaviors¹. A core characteristic of many psychiatric disorders is the abnormality of social behavior for instance, observed in major depressive disorder (MDD)²⁻⁵. An important symptom of MDD is the drastic alteration in social functioning, which can be expressed via the avoidance of social encounters^{6,7}, but also via maladaptive behaviors, such as aggressiveness or angriness^{8,9}. This can ultimately lead to social withdrawal from society, a risk factor for disease persistence and worsening of symptoms^{10,11}. Moreover, social dysfunction in MDD has been recently recognized as a semi-independent domain, which can even persist after the recovery from the core depressive symptoms^{11,12}.

The vulnerability towards MDD is determined by the combined effect of genetics, epigenetics, and the environment^{13,14}. Human genetic studies have identified different genomic variations that are involved in the susceptibility towards MDD, among others via several polymorphisms in the FKBP5 gene¹⁵⁻¹⁷. In addition, exposure to severe stress can have a detrimental effect on healthy living and ultimately is an important environmental risk factor for MDD vulnerability¹⁸⁻²⁰. The combination of severe stress exposure and genetic risk variants of the FKBP5 gene, encoding for the co-chaperone FK506 binding protein 51 (FKBP51), have been found to dysregulate the hypothalamic-pituitary-adrenal (HPA) axis and increase the risk for MDD vulnerability²¹⁻²⁵. FKBP51 is an important regulator of the HPA axis and stabilizes the GR-complex structure, thereby decreasing the binding to glucocorticoids and hampering the nuclear translocation of the GR complex²⁶. The region-specific role of FKBP51 and its impact on social functioning remains to be fully elucidated and comprehending their significance in MDD pathology is a crucial inquiry that must be addressed to enhance the mechanistic understanding of MDD. The genetic manipulation of FKBP5 is difficult to assess in humans but preclinical models using mice have shown to be a valuable tool to elucidate the behavioral and genetic influences of *Fkbp5* on MDD. Preclinical mouse studies have shown a high expression pattern of *Fkbp5* in, among others, the hypothalamus, amygdala, hippocampus, bed nucleus of stria terminalis, dorsal raphe nucleus (DRN), and locus coeruleus (LC)²⁷. The exposure to different stressors has shown a region-specific increase of *Fkbp5*, highlighting the importance of understanding the region-specific effects of *Fkbp5*²⁷⁻²⁹. The LC is an important region to further explore the specific influences of *Fkbp5*, since (1) this region shows high baseline *Fkbp5* expression²⁷, (2) plays a

crucial role in both stress response systems by initiating the sympathetic–adrenal–medullary (SAM) axis, and (3) has connections with the HPA axis via CRF neurons from several stress-related brain regions that can further innervate the LC^{30–32}.

The LC is strongly activated by stress exposure and has widespread noradrenergic projections throughout the brain, orchestrating large-scale network processes influencing many cognitive functions^{33,34}. In addition, the LC has pathway-specific influences. For example, the LC-induced norepinephrine (NE) release in the basolateral amygdala (BLA) increased anxiety-like behavior³⁵. Several studies have shown that the LC strongly responds to social stress exposure, via the general activation of the LC and adapting specific projection circuits^{36–38}, but the direct influence of the LC on social behavior remains to be uncovered.

The social behavioral construct entails a broad range of different behaviors³⁹, which are often too complicated, time-intensive, and repetitive to assess manually^{40–42}. Therefore, for decades social behavioral phenotyping has heavily utilized instrumental tasks that often rely on oversimplified behavioral constructs⁴². The advancement in automatically annotated behavioral assessment has allowed for high-throughput analysis using pose estimation^{43,44} and subsequently supervised classification⁴⁵. These tools fast-forward and simplify the behavioral analysis by using predefined behavioral constructs without losing the complexity of social behavior.

In the current study, we aimed to unravel the region-specific role of *Fkbp5* regulation on the underlying neurobiological mechanism and behavioral profile related to stress exposure. We employed the open-source Python package, “DeepOF” to enable a deep phenotyping strategy on the social behavioral profile to further elevate the social behavioral understanding related to stress exposure. We show a selectivity of the LC towards social stressors by cellular activation and *Fkbp5* expression patterns. Furthermore, the conditional knockout of the *Fkbp5* gene within the noradrenergic system shows an altered social behavioral profile response towards novel social stimuli only in male mice. In addition, we show that these mice, along with the social behavioral changes, show long-lasting changes in the noradrenergic signaling and related molecular pathways.

Materials and methods

Animals and housing

Wild-type adult male and female C57/Bl6N mice and the genetic mouse line *Fkbp5*^{Nat} (age between 2-3 months of age) were obtained from the in-house breeding facility of the Max Planck Institute of Psychiatry and used for breeding (F₀). The offspring (F₁) were used as experimental animals and were weaned at P25 in groups of maximum four animals with same-sex littermates. All experiments were performed during adulthood at the age of 3-5 months. All animals were housed in individually-ventilated cages (IVC; 30cm×16cm×16cm connected by a central airflow system: Tecniplast, IVC Green Line—GM500), while kept under standard housing conditions; 12h/12h light-dark cycle (lights on at 7 a.m.), temperature 23±1°C, humidity 55%. Food (Altromin 1324, Altromin GmbH, Germany) and water were available ad libitum. The experimental procedures were approved by the committee for the Care and Use of Laboratory Animals of the government of Upper Bavaria, Germany. All experiments were in accordance with the European Communities Council Directive 2010/63/EU.

Generation of the *Fkbp5*^{Nat} mouse line

The genetic mouse line *Fkbp5*^{Nat} is a conditional knock-out of the *Fkbp5* gene in the noradrenergic system, among others the LC. This was achieved via the generation (as performed by the knockout mouse project) of full knockout *Fkbp5*^{Frt/Frt} mice, which can re-express functional *Fkbp5* upon activation of the Flp recombinase. *Fkbp5*^{Frt/Frt} mice were then bred with Deleter-Flpe mice to create mice with a floxed *Fkbp5* gene; *Fkbp5*^{lox/lox} mice^{46,47}. Subsequently, the final conditional knock-out of *Fkbp5* was achieved via the crossing of *Fkbp5*^{lox/lox} mice with mice containing the Cre-cassette in a BAC-vector under the noradrenalin transporter gene (*Slc6a2*)⁴⁸, resulting in *Fkbp5*^{Nat} mice. The *Fkbp5*^{lox/lox} animals were used as wild-type (WT) control animals.

Chronic social defeat stress

A cohort of wild-type C57/Bl6N 2 months old male mice were either divided into the chronic social defeat stress (CSDS) protocol or were kept under normal housing conditions. The CSDS paradigm consisted of exposing the experimental C57Bl/6N mice to an aggressive CD1 mouse for 21 consecutive days, as previously described⁴⁹. The CD1 aggressor male mice were purchased from Janvier Labs (Germany) and were at least 16 weeks old. The CD1 aggressor mice were trained and selected on their aggression prior to the start of the experiment. Upon the start of the CSDS, the experimental mice were introduced daily to a new CD1 resident's territory, who subsequently attacked and forced the experimental mouse into

subordination. Defeat sessions lasted until the stress-exposed mouse received two bouts of attacks from the CD1 aggressor or at five minutes in the rare instances when two bouts were not achieved within this duration. Animal health was monitored throughout the experiment to ensure that minor injuries healed before the subsequent defeat session. Between daily defeats, stressed mice were housed in the resident's home cage but physically separated from the resident by a see-through, perforated mesh barrier, allowing sensory exposure to the CD1 aggressor mouse while preventing further attacks. The defeat time of day was randomized between 11 a.m. and 6 p.m. to avoid habituation and anticipatory behaviors in defeated mice. NS mice were single-housed in the same room as the stressed mice. All animals were handled daily and weighed every 3-4 days.

Acute stress exposure

The acute stress exposure consisted of either acute social defeat stress, or restraint stress. The acute stress was performed 24 hours after the end of the CSDS paradigm. The acute social defeat stress consisted of a single defeat event with a novel CD1 aggressor mouse, as described for the CSDS paradigm. The restraint stress consisted of the restraining in a 50ml falcon tube, that contained holes in the top and the lid to allow air ventilation and tail movement. Animals were restrained for 15min during which they were kept in their home-cage environment. Animals were sacrificed 4 hours after acute stress exposure, after which brains were removed and snap-frozen using 2-methyl butane (kept on dry ice) and stored at -80°C until further use.

In-situ* hybridization of *Fkbp5

The *Fkbp5* mRNA profile was determined using radio-active *in-situ* hybridization labeling as described previously²⁷. In brief, brains were sliced using a cryostat in 20 µm sagittal sections, which resulted in a series of LC (bregma: -5.34 to -5.80) and DRN (bregma: -4.36 to -4.84) slides that were thaw-mounted on Super Frost Plus Slides and stored at -20°C. The *in-situ* hybridization sections were removed from -20°C, left to dry at room temperature, fixated with 4% paraformaldehyde, and subsequently dehydrated using a series of increasing concentrations of ethanol. Then, the hybridization buffer was equally spread out over the different slides containing the radioactive ³⁵S-UTP-labeled *Fkbp5* riboprobe and incubated overnight at 55°C. On the next day, the sections were rinsed, incubated with RNase A, desalted, and dehydrated, after which the radioactive slides were exposed to Kodak Biomax MR films (Eastman Kodak Co., Rochester, NY) and developed after an exposure time of 12 days. Films were digitized and the regions of interest were identified using the mouse brain atlas (<https://mouse.brain-map.org/>). The expression

was determined by optical densitometry with the ImageJ software (NIH, Bethesda, MD, USA). The expression was averaged per brain region per animal and subtracted by the background signal of a nearby structure that did not express the *Fkbp5* gene.

C-Fos immunostaining

The C-Fos protein expression was determined using a DAB staining kit for immunohistochemistry (Abcam, USA; ab64261). Brains were sliced using a cryostat in 20 μm sagittal sections, which resulted in a series of LC slides that were thaw-mounted on Super Frost Plus Slides and stored at -20°C . Immunostaining was performed as described by the Abcam DAB staining protocol (ab64261). In short, the slides were fixated in 4% paraformaldehyde, then incubated with hydrogen peroxide (to block endogenous peroxidase), and then blocked in protein block to minimize unspecific binding. Then slides were incubated overnight at 4°C with the rabbit monoclonal C-Fos primary antibody (1:1000, ab222699), diluted in phosphate buffered saline (PBS) and 0.5% Bovine Serum Albumin. On the next day, the sections were incubated at room temperature for 10 min with the secondary antibody goat anti-polyvalent. Next, to amplify the signal, slides were incubated with streptavidin peroxidase. The DAB staining was performed by combining the DAB chromogen with the DAB substrate (1:50, respectively), which was then applied to the slides and exactly washed away after 3.5 min. Slides were dehydrated using a series of increasing concentrations of ethanol and cover slipped. Slides were then imaged using a slide scanner (Olympus, VS120-S6-W) on a 10x magnification using the bright field settings. Bilateral images were taken in a series of images of the LC, going from bregma -5.34 to -5.80. Ultimately, similar bregma images were taken for all animals and C-Fos puncta were counted for 2-3 images per animal per side.

Validation of the *Fkbp5*^{Nat} mouse line using RNAscope

Validation of the knock-out of *Fkbp5* in the LC was performed via an RNAscope *in-situ* hybridization study. Male mice were sacrificed under baseline conditions at 3-5 months of age. The brains were removed and snap-frozen using 2-methyl butane (kept on dry ice) and stored at -80°C until further use. Brains were sliced using a cryostat in 20 μm sagittal sections, which resulted in a series of LC slides that were thaw-mounted on Super Frost Plus Slides and stored at -20°C . The RNAscope staining procedures were used according to the manufacturer's protocol as previously described⁴⁷. The RNAscope fluorescent multiplex reagent kit (cat. no. 320850, Advanced Cell Diagnostics, Newark, CA, USA) was utilized for mRNA staining. The probes used for staining were *Fkbp5* (Probe: Mm-Fkbp5-C1), and Tyrosine hydroxylase (TH) (Probe: Mm-TH-C2). Slides were then imaged using a ZEISS confocal microscope on a 40x magnification using the

fluorescent channel. Bilateral images were taken in a series of images of the LC, going from bregma -5.34 to -5.52. All images were acquired using the same settings for laser power, detector gain, and amplifier offset. *Fkbp5* mRNA expression was analyzed using ImageJ with the experimenter blinded to the genotype of the animals and was counted manually. Ultimately, similar bregma images were taken for all animals and *Fkbp5* puncta were counted within TH-positive cells for 2-3 images per animal per side. *Fkbp5* negative TH cells were accounted for when the cell had less than five *Fkbp5* puncta. Ultimately, a calculation per animal was made for the percentage of *Fkbp5-positive* TH cells compared to the total amount of TH cells.

***Fkbp5*^{Nat} adult testing**

At 3 months of age, a cohort of both males and females were tested in the social interaction task. In addition, separate cohorts of male *Fkbp5*^{Nat} were used to do microdissection, proteomics, and microdialysis experiments. The behavioral tests were performed between 8 a.m. and 11 a.m. in the same room as the housing facility.

Social interaction task using DeepOF analysis

The social interaction task was performed in a round open field arena (diameter of 38cm) using sawdust material on the bottom, as previously described by ⁴⁵. The experimental animal was placed in the open field arena and could freely explore for 10 min, after which an unfamiliar young CD1 (4–6 weeks old) social conspecific was placed in the same arena and both were allowed free exploration of the arena and each other for 10min. The data was recorded using the DFK37BUX250 imaging source (Germany) cameras with camera lenses from Stoelting, Ireland (Item nr. 60528). The IC capture software (version 2.5.1547.4007) from imaging source was used to obtain the videos and further analysis was performed with DeepLabCut version 2.2b7 (single animal mode for pose estimation ^{43,44}, and subsequently, DeepOF module version 0.1.6 ⁴⁵ for supervised behavioral analysis of six individualistic behaviors during the open field; wall-climbing, digging, huddling, look-around, sniffing, and speed (locomotion), and all during the social interaction task, including an additional five social behaviors; nose-to-nose, Side-by-side, Side-reverse-side, nose-to-tail, and Nose-to-body. Data were analyzed for the total 10 min of both tasks and in time bins of 2.5 min.

Microdissection

The total concentration of NE was measured from microdissected fresh frozen brain tissue in both WT and KO mice comparing no interaction (NI) and social interaction conditions. At the start of the experiment, mice were randomly divided into either the NI or SI condition. The NI animals were left

undisturbed in their home-cage, whereas SI animals were exposed to an unfamiliar young CD1 (4–6 weeks old) mouse for 10min within their home-cage. After the 10min SI exposure, all animals were directly sacrificed and brains were removed and snap-frozen using 2-methyl butane (kept on dry ice) and stored at -80°C until further use. Then, the brains were sectioned using a VT1200/S Leica vibratome on 250µm thick slices by the different brain regions; 2 slices in the medial prefrontal cortex (mPFC) (bregma: 1.94 to 1.54), 3 slices in the basolateral amygdala (BLA) (bregma: -1.34 to -1.94), 3 slices in the dorsal hippocampus (dHipp) (bregma: -1.70 to -2.18), and 3 slices in the ventral hippocampus (bregma: -3.08 to -3.52). The sliced tissue was directly punched within the vibratome using a sample corer (diameter 1 mm) and stored in 1.5 mL Safe-lock Eppendorf tubes on dry ice and subsequently stored at -80°C until further use. The measurement of NE was carried out by reverse-phase liquid chromatography with electrochemical detection as described in ⁵⁰. The values obtained were expressed as nanograms per milligram wet tissue and were logarithmically transformed for calculation of linearity of regression, standard error of the regression coefficients, and significance of differences between regression coefficients.

Proteomics

The proteomics analysis was performed from microdissected fresh frozen brain tissue in both WT and KO mice under baseline conditions. The same extraction protocol was performed as described at the microdissection, but only the BLA was punched for proteomic analysis. BLA tissue punches were homogenized in ice-cold T-PER™ tissue protein extraction reagent (ThermoFisher Scientific, 78510) freshly supplemented with protease inhibitor cocktail tablets (Roche, 05892791001) and phosphatase inhibitor cocktail tablets (Roche, 04906837001). Protein extracts were centrifuged at 10,000 x g for 10 minutes to pellet cell/tissue debris. Subsequently, protein concentration was adjusted to 2µg/µl using the Pierce™ BCA protein assay kit (ThermoFisher Scientific, 23225). For mass spectrometry measurements, in-solution samples were sent to the Max Planck Institute of Biochemistry Core Facility, Mass Spectrometry Lab.

Microdialysis

The microdialysis experiment consisted of the measurement of NE and its metabolite 3-Methoxy-4-hydroxyphenylglycol (MHPG) unilaterally in the BLA during baseline and social interaction in WT and KO animals. The microdialysis workflow was performed as described previously ⁵¹. In brief, male mice at 3 months of age were anesthetized with isoflurane and fixated in a stereotactic apparatus. Then microdialysis guide cannula was inserted in the right BLA (bregma: AP -1.35mm, ML 3.3 mm, and DV

4.4mm). After surgery, animals were treated with meloxicam for three days and were allowed to recover for a minimum of 7 days in the home-cage environment.

The microdialysis system contained a syringe pump (Harvard Apparatus, USA) that was connected via FEP tubing (CMA, Cat. N. 8409501) to a dual-channel liquid swivel (Microbiotech Se) and could then be connected to the probe via FEB tubing. The perfusion liquid consisted of artificial CSF, which consisted of a solution of NaCl (0.86%), KCl (0.020%), MgCl₂·6H₂O (0.024%), and CaCl₂·2H₂O (0.018%) in HPLC water with a PH of 7.4. Then, the microdialysis probe (CMA 7 Probe 1 mm membrane, 6 kDa; Cat.N. 000082) was connected to the running microdialysis system using a constant flow rate of 1.5µl/min. Any air bubbles in the probe were removed, after which the probe was inserted by hand into the guide cannula of the animal at least 20 hours before the start of the experiment. Animals stayed in specific microdialysis cages containing a standard amount of sawdust material and nesting material (16 cm length x 16 cm width x 32 cm height). The experiment started with the collection of baseline samples, after which an unfamiliar young CD1 (4–6 weeks old) mouse was introduced into the microdialysis cages, resulting in the comparison of baseline (NI) and social interaction (SI) samples. The measurement of NE and MHPG out of the microdialysates was performed using HPLC with electrical detection, as previously described by ⁵¹. Quantification was performed using external standard calibration (0.1–5 nM).

Statistics

Statistical analyses and graphs were made using Rstudio (with R 4.1.1). All animals were used for statistical analysis unless stated otherwise. Statistical outlier tests were performed using boxplot analysis, in which values above quartile (Q) $3 + 1.5 \times$ interquartile range (IQR, calculated as $Q3 - Q1$) or below $Q1 - 1.5 \times$ IQR were considered outliers. This led to the exclusion of 9 animals in the microdialysis experiment, 4 WT and 5 KO mice. Data were tested on the corresponding statistical assumptions, which included the Shapiro-Wilk test for normality and Levene's test for heteroscedasticity. If assumptions were violated the data were analyzed using non-parametric variants of the test. The two group comparisons were analyzed using the independent samples t-test (T) as a parametric option, Welch's test (We), if data was normalized but heteroscedastic, or the Wilcoxon test (Wx) as a non-parametric option. The chronic and acute stress exposure data using a six-group comparison was analyzed using the one-way ANOVA (parametric) or Kruskal-Wallis test (non-parametric), and further posthoc analysis was performed using the BH test (parametric) or the Wilcoxon test (non-parametric). The metabolite measurement data in the microdissection and microdialysis experiments were analyzed using a two-way ANOVA with the genotype (WT or KO) as within-subject factor and the social interaction (baseline vs. social interaction) as a between-

subject factor. P-values were adjusted for multiple testing using the BH method. The bar graphs are presented as mean \pm standard error of the mean (SEM). Data were considered significant at $p < 0.05$ (*), and further significance was represented as $p < 0.01$ (**), $p < 0.001$ (***), and $p < 0.0001$ (****).

Results

The *Fkbp5* gene shows a social stress-specific upregulation in the LC

The expression patterns of *Fkbp5* were examined in the LC and DRN brain regions across various chronic and acute stress paradigms. Stress-naïve mice were sacrificed under baseline conditions (NS) or exposed to an acute stressor before sacrifice, which was either acute social defeat stress (ASDS), or restraint stress (Res). Another group of animals underwent 21 days of chronic social defeat stress (CSDS) followed by sacrifice 24 hours later under baseline conditions (CSDS) or following acute stress exposure, which was either acute social defeat stress (CSDS+ASDS), or restraint stress (CSDS+Res), as depicted in Figure 1A.

In mice without a history of chronic stress, acute social defeat stress led to a significant increase in *Fkbp5* mRNA levels compared to stress-naïve and acutely restrained mice in both the LC (Figure 1B-C) and DRN (Supplemental Figure 1A-B) regions. A similar effect was observed in mice with a history of chronic stress in the DRN region, where mice exposed to CSDS plus acute social defeat stress exhibited elevated *Fkbp5* mRNA levels compared to those in CSDS baseline conditions and CSDS plus acutely restrained conditions (Supplemental Figure 1A-B). However, the LC region displayed a different response after chronic stress, as mice exposed to CSDS plus acute social defeat stress did not show elevated *Fkbp5* mRNA levels compared to those in CSDS baseline conditions and CSDS plus acutely restrained conditions. Furthermore, the baseline acute social defeat stress mice demonstrated significantly higher *Fkbp5* mRNA levels compared to animals with a chronic stress history plus acute social defeat stress mice (Figure 1B-C).

The expression levels of the C-Fos protein in the LC exhibited a similar pattern to *Fkbp5* expression under acute stress conditions. Exposure to acute social defeat stress significantly upregulated C-Fos expression compared to stress-naïve and acutely restrained mice (Figure 1D-E). In contrast to the *Fkbp5* results, a history of CSDS exposure did not result in altered C-Fos expression pattern between mice exposed to CSDS plus acute social defeat stress and those in CSDS baseline conditions or CSDS plus acutely restrained conditions (Figure 1D-E). Additionally, the baseline acute social defeat stress mice did not exhibit altered C-Fos expression compared to animals with a chronic stress history plus acute social defeat stress (Figure 1D-E).

The conditional knock-out of *Fkbp5* alters the social behavioral profile in male mice

Considering the high expression of *Fkbp5* in the LC and its specific *Fkbp5* responsiveness to social stressors, a genetic conditional *Fkbp5* knockout line was generated using the *Nat* promoter. This conditional knockout line was employed to investigate the precise effects of *Fkbp5* knockout on the noradrenergic system, including the LC (Figure 2A). Through RNAscope co-expression analysis of tyrosine hydroxylase (TH), a marker for identifying noradrenergic cells, and *Fkbp5*, a significant reduction in the percentage of *Fkbp5*-positive TH cells within the LC was observed in *Fkbp5*^{Nat} (KO) mice compared to *Fkbp5*^{lox/lox} (WT) mice (Figure 2B-C).

Subsequently, the social behavioral profile was examined during the social interaction task in both male and female mice, utilizing four distinct time bins of 2.5 minutes each. Principal Component Analysis (PCA) was employed to explore the social behavior profile across different time bins, including both sexes and genotypes. The PCA results indicated a disparity between the first 2.5-minute bin and all subsequent time bins, suggesting that the initial 2.5 minutes were particularly noteworthy for social behavioral phenotyping (Supplemental Figure 2A-B). To assess the differences between sexes, a PCA was conducted using the first 2.5-minute social interaction bin, regardless of genotype. Notably, a significant distinction was observed between female and male mice, suggesting a contrasting social behavioral profile between the sexes. Consequently, subsequent analyses were performed separately for each sex (Supplemental Figure 2C-D).

Separate PCAs were conducted for female and male social interaction data to compare genotypes. In the female PCA, no significant differences were found between genotypes (Figure 2D). However, in the male PCA, a significant difference was observed (Figure 2E). Further examination of the male PCA data was performed via the exploration of the top contributing behaviors in PC1, determined by the corresponding rotated loading scores (Figure 2F). The top five contributing behaviors displayed potentially relevant patterns for identifying genotype effects, while other behaviors within the top 10 either contributed to the CD1 animal ("W-" behaviors) or had low rotated loading scores. Further analysis of the top contributing behaviors in the male 2.5-minute social interaction data revealed a significant increase in KO mice compared to WT mice in the expression of B-nose-to-tail, B-nose-to-body, and B-following behaviors, whereas no genotype difference was observed in B-look-around and B-speed behaviors (Figure 2G-K). To investigate the specificity of acute effects of social exposure, male mice were examined for their social behavioral profile following chronic social defeat stress (CSDS) exposure. The PCA results did not indicate

any alterations between genotypes, suggesting a specific effect of acute social exposure in male KO mice (Supplemental Figure 2E-F).

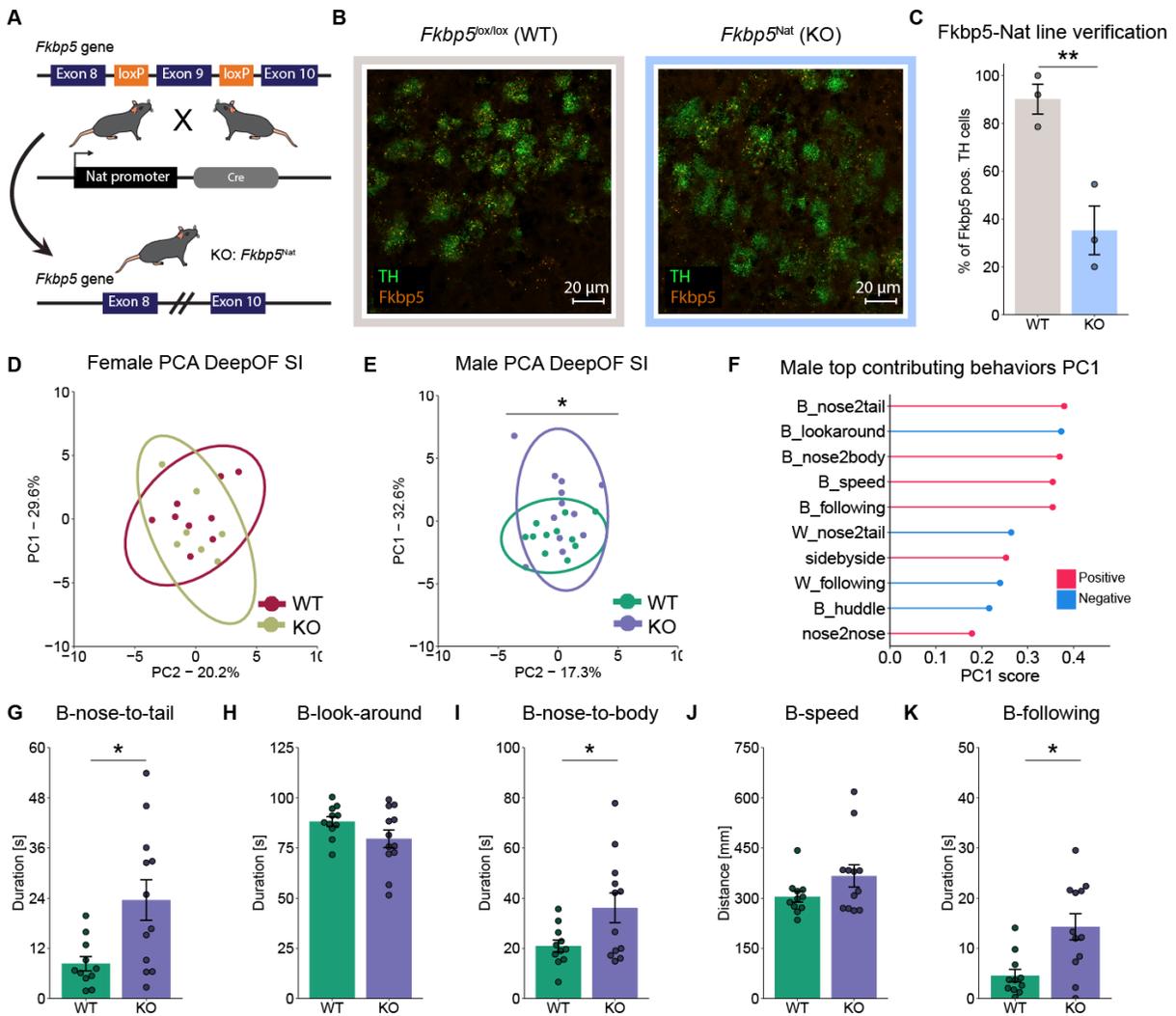


Figure 2 FKBP51-Nat males show an altered social behavioral profile. A) Schematic overview of the generation of *Fkbp5* knock-out in the noradrenergic cells. **B)** Representative RNAscope confocal images of *Fkbp5* mRNA expression in TH neurons within the LC. **C)** RNAscope *Fkbp5* quantification in the TH neurons revealed a significant reduction of the percentage of *Fkbp5* positive TH cells in KO animals (Two-tailed independent samples-t-test: $T(4)=4.6$, $p=0.01$). **D)** The female SI 2.5 min time bin PCA showed no difference in the PC1 eigenvalues between conditions. The PCA data consisted of all the SI DeepOF behavioral classifiers. Two-tailed independent samples t-test: $T(14)=-0.72$, $p=0.49$. **E)** The male SI 2.5 min time bin PCA showed a significant difference in the PC1 eigenvalues between conditions. Welch test: $We(14.96)=-2.32$, $p=0.035$. **F)** The top contributing behaviors for males in the SI 2.5 min time bin PC1 using the corresponding rotated loading scores. The top five behaviors were listed as potentially relevant behaviors for identifying genotype effect (B-nose-to-tail (0.38), B-look-around (-0.37), B-nose-to-body (0.37), B-speed (0.35), B-following (0.35). "B-" indicates C57Bl/6N behaviors and "W-" indicates CD1 behaviors. **G)** The 2.5 min duration of B-nose-to-tail.

Wilcoxon test: $W_x=27$, $p=0.016$. **H**) The 2.5 min duration of B-look-around. Welch test: $W_e(16,90)=1.71$, $p=0.11$. **I**) The 2.5 min duration of B-nose-to-body. Wilcoxon test: $W_x=38$, $p=0.091$. **J**) The 2.5 min duration of B-speed. Wilcoxon test: $W_x=47$, $p=0.26$. **K**) The 2.5 min duration of B-following. Wilcoxon test: $W_x=25$, $p=0.011$. In panel C; $n=3$ for WT and $n=3$ for KO. In panel D; $n=9$ for WT and $n=7$ for KO. In panels E-K; $n=11$ for WT and $n=12$ for KO.

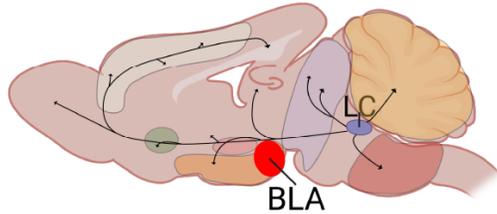
Noradrenergic signaling is altered in *Fkbp5*^{Nat} within the BLA

The impact of conditional knockout of *Fkbp5* within the noradrenergic system was investigated using *Fkbp5*^{Nat} (KO) mice. This was examined by the total quantity of NE in various projection regions of the LC following baseline conditions (NI) and social interaction (SI) (Supplemental Figure 3A). The total amount of NE was not significantly altered by genotype or social interaction exposure in the mPFC, dHipp, and vHipp (Supplemental figure 3 B-D). On the contrary, the BLA region (Figure 3A) showed a significant increase in total NE concentration after social interaction exposure compared to baseline in WT, but not in KO mice (Figure 3B). To gain further insight into the BLA region-specific protein profile under baseline conditions, a proteomics analysis was conducted, comparing WT and KO mice. The Panther GO TERM pathway analysis, employing the top 100 up- and downregulated proteins, exhibited a significant and substantial fold enrichment for the dopaminergic and noradrenergic signaling pathway (GO:0007191) and the neurotransmitter metabolic process (GO:0042133). Considering the knockout of *Fkbp5* within the noradrenergic system, further investigation focused on the pathway associated with the noradrenergic synapse. The proteomics analysis of the noradrenergic synapse unveiled several downregulated proteins in KO mice compared to WT mice, including tyrosine hydroxylase (TH), AC5, and DARPP32, alongside one upregulated protein called MAPK8. Notably, only the protein tyrosine hydroxylase exhibited a significant alteration (Figure 3D).

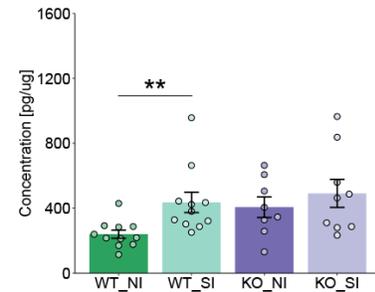
Given the differences unveiled in the molecular pathway associated with noradrenergic signaling in the BLA between KO and WT mice, our subsequent question focused on scrutinizing norepinephrine (NE) release in the BLA. To accomplish this, we conducted a microdialysis experiment, wherein we examined NE release across genotypes (WT vs KO) and social interaction history (baseline (NI) vs social interaction (SI)). Following social interaction, the turnover rate of MHPG/NE in the BLA showed a significant reduction compared to the baseline levels in WT mice but not in KO mice (Figure 3E). A similar trend was observed for NE release, as indicated by the elevation in NE concentration after social interaction compared to

baseline specifically in WT mice (Supplemental Figure 3E). No notable differences were detected in MHPG release based on genotype or exposure to social interaction (Supplemental Figure 3F).

A Schematic punched project regions LC

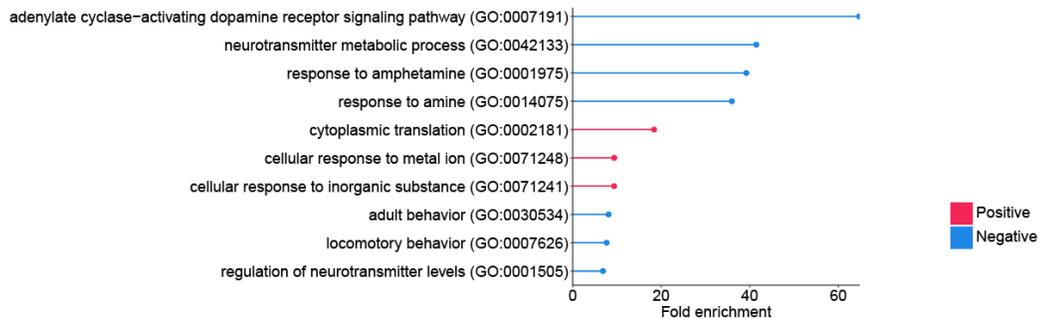


B Microdissection BLA - NE



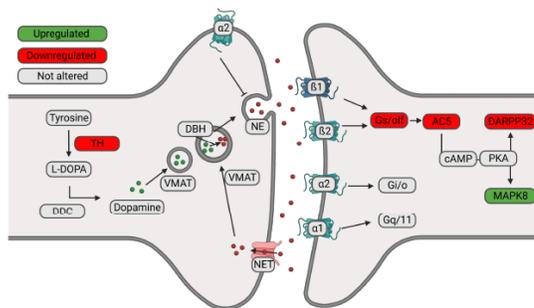
C

Proteomics - Panther GO TERM



D

Proteomics - pathway analysis - the noradrenergic synapse



E Microdialysis BLA - MHPG/NE

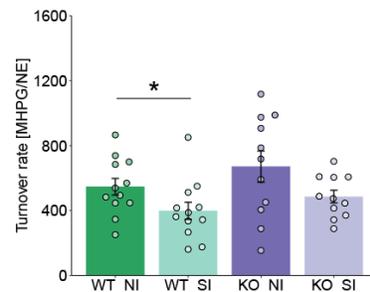


Figure 3 Noradrenergic signaling is altered in *Fkbp5*^{Nat} within the BLA. **A)** Schematic overview of the LC projections, in which the BLA was punched. **B)** Microdissection measuring total NE levels in the BLA. Posthoc Kruskal-Wallis revealed that WT social interaction (WT_SI) significantly increased NE concentrations in the BLA compared to WT no interaction (WT_NI), $F(1)=10.14$, $p=0.0029$, which was not altered in the *Fkbp5*^{Nat} background, comparing KO_NI with KO_SI, $p=1$. **C)** The proteomics GO TERM analysis between WT and KO under baseline by Panther identified interesting molecular pathways within biological process, in particular interest were “GO:0007191”, $FDR=64.70$, $p=2.64 \times 10^{-5}$, and “GO:0042133”, $FDR=41.47$, $p=2.93 \times 10^{-7}$. **D)** The pathways analysis related to the noradrenergic synapse is highlighted with multiple proteins altered. The proteins downregulated were tyrosine hydroxylase (TH); $D=-0.70$, $p=0.049$, Gs/olf; $D=-0.41$, $p=0.13$, ACS; $D=-0.45$, $p=0.14$, DARPP32; $D=-0.52$, $p=0.054$ and the upregulated protein was: MAPK8 $D=0.72$, $p=0.11$. **E)** Microdialysis measuring turnover rate of MHPG/NE release in the BLA. Posthoc BH revealed that WT social interaction (WT_SI) has significantly lowered NE/MHPG turnover rate in the BLA compared

to WT no interaction (WT_NI), $F(1,22)=4.01$, $p=0.05$, which was not altered in the *Fkbp5*^{Nat} background, comparing KO_NI with KO_SI, $p=0.088$. Two-way ANOVA: social interaction; $F(1,42)=7.07$, $p=0.011$, genotype; $F(1,42)=0.287$, $p=0.098$, social interaction*genotype; $F(1,42)=0.085$, $p=0.77$. In panel B; $n=11$ for WT_NI, $n=11$ for WT_SI, $n=8$ for KO_NI, and $n=9$ for KO_SI. In panels C-D; $n=9$ for WT and $n=12$ for KO. In panel E; $n=12$ for WT and $n=11$ for KO (within-subject analysis on between baseline and social interaction).

Discussion

The LC is a crucial node within the stress response system that integrates input from many different stress-related neural circuits and has widespread noradrenergic projections through the entire neuraxis, orchestrating large-scale network processes but also pathway-specific networks³⁵. The combination of severe stress exposure and genetic risk variants of the FKBP5 gene have been found to increase the risk for MDD pathology^{21–23}, but the region-specific role of the FKBP5 gene and its impact on social functioning remains to be fully elucidated. The current study shows the selectivity of the LC towards social stressors by cellular activation and *Fkbp5* expression patterns in mice. Furthermore, a conditional knockout of *Fkbp5* within the noradrenergic system shows an altered social behavioral profile in male mice using a deep phenotyping strategy, which has long-lasting consequences on noradrenergic signaling and the related molecular pathways.

***Fkbp5* expression in the LC reacts to stress in a highly selective way**

Previous studies have provided evidence supporting the brain region-specific upregulation of *Fkbp5* mRNA, which is contingent upon the specific type of stressor employed^{27–29}. These findings emphasize the significance of considering the context and nature of the stressor in the regulation of *Fkbp5* expression. Specifically, acute restraint stress has been shown to upregulate *Fkbp5* in the paraventricular nucleus of the hypothalamus (PVN) and central amygdala, while acute food deprivation leads to *Fkbp5* upregulation in the hippocampus²⁷. Additionally, chronic non-social stress exposure increases *Fkbp5* mRNA levels in the hippocampus and prefrontal cortex but not in the hypothalamus²⁸.

The present study highlights the importance of stress-specific selectivity in the regulation of *Fkbp5* mRNA, as evidenced by the highly selective upregulation of *Fkbp5* in the LC following acute social stress but not chronic social stress or restraint stress. Consistent with previous research demonstrating brain-region-specific upregulation of *Fkbp5*²⁷, we observed a distinct pattern of *Fkbp5* upregulation in the DRN. Both

acute social stress and acute social stress in mice with a history of chronic social stress led to an upregulation of *Fkbp5* in the DRN. This study demonstrates that the regulation of *Fkbp5* expression in response to social stress varies across different brain regions. Previous research has shown that acute stress exposure differentially alters LC excitability over time, with increased LC excitability observed one week after stress exposure compared to immediately after the stress exposure³⁶. The specific increase in *Fkbp5* mRNA levels may be attributed to heightened cellular activity in the LC region during acute social stress. To exclude the possibility that the observed differential expression pattern of *Fkbp5* is merely a reflection of cellular activity, we examined the expression patterns of C-Fos across the different stress groups. C-Fos expression serves as an indicator of cellular activity within the LC region, as previously investigated in various stress paradigms and brain regions^{52–55}. Interestingly, C-Fos protein expression levels in the LC displayed a similar following acute stress, independent of chronic stress history and thereby distinct from the differential *Fkbp5* expression pattern. Acute social stress induced a robust C-Fos upregulation, which was not altered compared to acute social stress after a history of chronic social defeat stress. This finding aligns with Reyes et al., 2019⁵⁶, who demonstrated that a shorter 5-day CSDS protocol exhibited a much lower C-Fos count compared to acute social stress. Therefore, it can be concluded that the differential expression pattern of *Fkbp5* in the LC based on stress history is not merely a reflection of regional activity. The specific and robust upregulation of *Fkbp5*, particularly following acute social stress exposure, is dampened after chronic social stress, suggesting the significance of the social nature of the stressor in the regulation of this risk gene in the LC.

***Fkbp5*^{Nat} alters the social behavioral profile and noradrenergic signaling in male mice**

The context of social stress exposure appears to be particularly relevant for *Fkbp5* expression in the LC, prompting the generation of a conditional *Fkbp5* knockout line under the Nat promoter (*Fkbp5*^{Nat}, KO) to investigate the specific effects of *Fkbp5* knockout on the noradrenergic system, including the LC. The social behavioral profile was investigated using advanced automated behavioral assessment tools based on predefined behavioral constructs, eliminating the need for labor-intensive manual labeling or the loss of complexity of the social behavioral construct. The results indicate that the overall social behavioral profile was specifically altered in male KO mice, while female KO mice did not show such alterations. Highly affected behaviors included increased following, nose-to-tail, and nose-to-body contact durations between the experimental mouse and the social conspecific, indicating an increase in pro-social behavior in male KO mice. Moreover, no changes in the social behavioral profile were observed in male mice based on genotype after chronic social stress exposure. Thus, the alterations in social behavior align with the

pattern of *Fkbp5* regulation in the LC, as only acute exposure to a social event led to changes in the social behavioral profile and *Fkbp5* expression pattern. A recent study by Nold et al. 2022⁵⁷ investigated the impact of human *FKBP5* single nucleotide polymorphism (SNP) variants on social avoidance behavior in mice using a three-chamber social arena task. They found a sex-specific effect, with only female *FKBP5* SNP resilience-associated (C/G) mice showing a preference for social interaction compared to *FKBP5* SNP risk-associated (A/T) mice⁵⁷. This study highlights the role of sex-dependent social behavioral effects related to *Fkbp5* manipulation. However, the use of the three-chamber social arena task limits the assessment of social behavioral read-outs and does not fully capture the same type of ethologically relevant social behaviors observed in the current study using DeepOF⁴².

To understand the mechanistic changes caused by *Fkbp5* knockout in the noradrenergic system, the total concentration of NE was examined in various LC projection regions. Interestingly, only the BLA region exhibited a significant increase in NE concentration following exposure to social interaction compared to baseline, specifically in WT mice. In contrast, KO mice did not show this effect. Panther GO TERM pathway proteomics analysis confirmed the impact of *Fkbp5*^{Nat} on noradrenergic synapse in the BLA, suggesting specific pathway changes within the synapse. To further investigate the alterations in the NE system in the BLA, a microdialysis experiment was conducted to assess NE release across genotypes under baseline and social interaction conditions. The turnover of MHPG/NE was decreased, and NE release was increased by social interaction exclusively in WT mice. These findings indicate a deficiency in the response of NE signaling in KO mice specifically in the BLA after exposure to a social stimulus. Therefore, it can be concluded that *Fkbp5* knockout in the noradrenergic system alters noradrenaline signaling and molecular pathways associated with the noradrenergic synapse, particularly in the BLA. Previous research has established the crucial role of the noradrenergic pathway in the BLA, where it is heavily involved in emotional memory consolidation⁵⁸⁻⁶⁰. Additionally, McCall et al. 2017³⁵ demonstrated that specific stimulation of LC-NE fibers in the BLA leads to NE release, alters BLA activity patterns, and increases anxiety-like behavior. These studies demonstrate the significance of the LC-BLA pathway in modulating behavior and the physiological properties of the BLA region via altered activity patterns. The current study sheds light on the role of *Fkbp5* regulation within the noradrenergic system in modulating the signaling pathways of NE specifically within the BLA. The findings suggest that *Fkbp5* plays a role in regulating the noradrenergic signaling within the BLA and further contribute to our understanding of the molecular mechanisms underlying social behavior.

Conclusion

In summary, the current study provides several key findings. First, it demonstrates that *Fkbp5* expression in the LC is selectively regulated in response to different stressors, with acute social stress specifically leading to *Fkbp5* upregulation. This highlights the importance of considering the social nature of stressors when studying *Fkbp5* regulation in the LC. Second, the study shows that altered *Fkbp5* regulation in the noradrenergic system has a sex-dependent impact on the social behavioral profile towards a novel social conspecific. This suggests that *Fkbp5* plays a role in modulating social behavior, with the effects varying between male and female mice. Lastly, the study uncovers the role of *Fkbp5* in modulating noradrenergic signaling within the BLA, particularly in response to a novel social stimulus. These findings contribute to our understanding of how genetic and environmental factors, such as FKBP5 and stress exposure, interact to influence psychiatric risk and social behavior-related symptoms. Overall, this research provides valuable insights into the complex mechanisms underlying social behavior and the potential implications for psychiatric disorders.

Acknowledgements

The authors thank Daniela Harbich, Bianca Schmid, and Cornelia Flachskamm for their technical assistance, Marius Fuchs and Prof. Dr. Jochen Klein for the assistance with the microdialysis experiments, and the DeepLabCut development team for creating and maintaining the DeepLabCut software.

Author contributions

JB and MVS conceived the study. JB performed the experiments, TB, SC, EA, LMB, LD, SN, HY, SM, VK, MS, MA, and NG assisted with the experiments. JD provided the genetic mouse line. JB analyzed the data and was assisted by LM and BMM.

Funding

This study is supported by the “Kids2Health” grant of the Federal Ministry of Education and Research [01GL1743C].

References

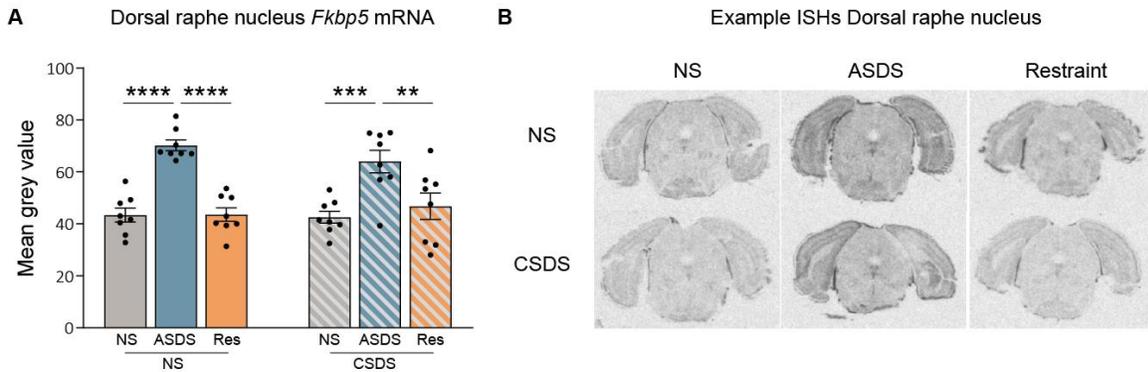
1. Faure, P., Fayad, S. L., Solié, C. & Reynolds, L. M. Social Determinants of Inter-Individual Variability and Vulnerability: The Role of Dopamine. *Front Behav Neurosci* **16**, (2022).
2. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. Am Psychiatric Assoc vol. 21 (2013).
3. Nietlisbach, G. & Maercker, A. Social Cognition and Interpersonal Impairments in Trauma Survivors with PTSD. *J Aggress Maltreat Trauma* **18**, 382–402 (2009).
4. Dodell-Feder, D., Tully, L. M. & Hooker, C. I. Social impairment in schizophrenia. *Curr Opin Psychiatry* **28**, 236–242 (2015).
5. Katz, S. J., Conway, C. C., Hammen, C. L., Brennan, P. A. & Najman, J. M. Childhood Social Withdrawal, Interpersonal Impairment, and Young Adult Depression: A Mediational Model. *J Abnorm Child Psychol* **39**, 1227–1238 (2011).
6. Fernández-Theoduloz, G. *et al.* Social avoidance in depression: A study using a social decision-making task. *J Abnorm Psychol* **128**, 234–244 (2019).
7. Ottenbreit, N. D., Dobson, K. S. & Quigley, L. An examination of avoidance in major depression in comparison to social anxiety disorder. *Behaviour Research and Therapy* **56**, 82–90 (2014).
8. Riley, W. T., Treiber, F. A. & Woods, M. G. Anger and Hostility in Depression. *J Nerv Ment Dis* **177**, (1989).
9. Painuly, N., Sharan, P. & Mattoo, S. K. Relationship of anger and anger attacks with depression. *Eur Arch Psychiatry Clin Neurosci* **255**, 215–222 (2005).
10. Kupferberg, A., Bicks, L. & Hasler, G. Social functioning in major depressive disorder. *Neurosci Biobehav Rev* **69**, 313–332 (2016).
11. Porcelli, S. *et al.* Social brain, social dysfunction and social withdrawal. *Neurosci Biobehav Rev* **97**, 10–33 (2019).
12. Gur, R. C. & Gur, R. E. Social cognition as an RDoC domain. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* **171**, 132–141 (2016).
13. de Kloet, E. R., Joëls, M. & Holsboer, F. Stress and the brain: from adaptation to disease. *Nat Rev Neurosci* **6**, 463–475 (2005).
14. Binder, E. B. The role of FKBP5, a co-chaperone of the glucocorticoid receptor in the pathogenesis and therapy of affective and anxiety disorders. *Psychoneuroendocrinology* **34**, 186–195 (2009).
15. Binder, E. B. *et al.* Polymorphisms in FKBP5 are associated with increased recurrence of depressive episodes and rapid response to antidepressant treatment. *Nat Genet* **36**, 1319–1325 (2004).
16. Fan, B. *et al.* Association of FKBP5 gene variants with depression susceptibility: A comprehensive meta-analysis. *Asia-Pacific Psychiatry* **13**, (2021).

17. Ising, M. *et al.* FKBP5 Gene Expression Predicts Antidepressant Treatment Outcome in Depression. *Int J Mol Sci* **20**, 485 (2019).
18. Heim, C. & Nemeroff, C. B. The role of childhood trauma in the neurobiology of mood and anxiety disorders: preclinical and clinical studies. *Biol Psychiatry* **49**, 1023–1039 (2001).
19. Kessler, R. C. *et al.* Childhood adversities and adult psychopathology in the WHO World Mental Health Surveys. *British Journal of Psychiatry* **197**, 378–385 (2010).
20. Nemeroff, C. B. Paradise Lost: The Neurobiological and Clinical Consequences of Child Abuse and Neglect. *Neuron* **89**, 892–909 (2016).
21. Ferrer, A. *et al.* FKBP5 polymorphisms and hypothalamic-pituitary-adrenal axis negative feedback in major depression and obsessive-compulsive disorder. *J Psychiatr Res* **104**, 227–234 (2018).
22. Klengel, T. *et al.* Allele-specific FKBP5 DNA demethylation mediates gene – childhood trauma interactions. *Nat Neurosci* **16**, (2013).
23. Rao, S. *et al.* Common variants in FKBP5 gene and major depressive disorder (MDD) susceptibility: a comprehensive meta-analysis. *Sci Rep* **6**, 32687 (2016).
24. Binder, E. B. *et al.* Association of FKBP5 polymorphisms and childhood abuse with risk of posttraumatic stress disorder symptoms in adults. *Journal of American Medical Association* **299**, 1291–1305 (2008).
25. Wang, Q., Shelton, R. C. & Dwivedi, Y. Interaction between early-life stress and FKBP5 gene variants in major depressive disorder and post-traumatic stress disorder: A systematic review and meta-analysis. *J Affect Disord* **225**, 422–428 (2018).
26. Baker, J., Ozsan, I., Rodriguez Ospina, S., Gulick, D. & Blair, L. Hsp90 Heterocomplexes Regulate Steroid Hormone Receptors: From Stress Response to Psychiatric Disease. *Int J Mol Sci* **20**, 79 (2018).
27. Scharf, S. H., Liebl, C., Binder, E. B., Schmidt, M. V. & Müller, M. B. Expression and regulation of the Fkbp5 gene in the adult mouse brain. *PLoS One* **6**, 1–10 (2011).
28. Guidotti, G. *et al.* Glucocorticoid receptor and fkbp5 expression is altered following exposure to chronic stress: Modulation by antidepressant treatment. *Neuropsychopharmacology* **38**, 616–627 (2013).
29. Santarelli, S. *et al.* An adverse early life environment can enhance stress resilience in adulthood. *Psychoneuroendocrinology* **78**, 213–221 (2017).
30. Valentino, R. J., Foote, S. L. & Aston-Jones, G. Corticotropin-releasing factor activates noradrenergic neurons of the locus coeruleus. *Brain Res* **270**, 363–367 (1983).
31. McCall, J. G. *et al.* CRH Engagement of the Locus Coeruleus Noradrenergic System Mediates Stress-Induced Anxiety. *Neuron* **87**, 605–620 (2015).
32. Valentino, R. J. & Van Bockstaele, E. Convergent regulation of locus coeruleus activity as an adaptive response to stress. *Eur J Pharmacol* **583**, 194–203 (2008).

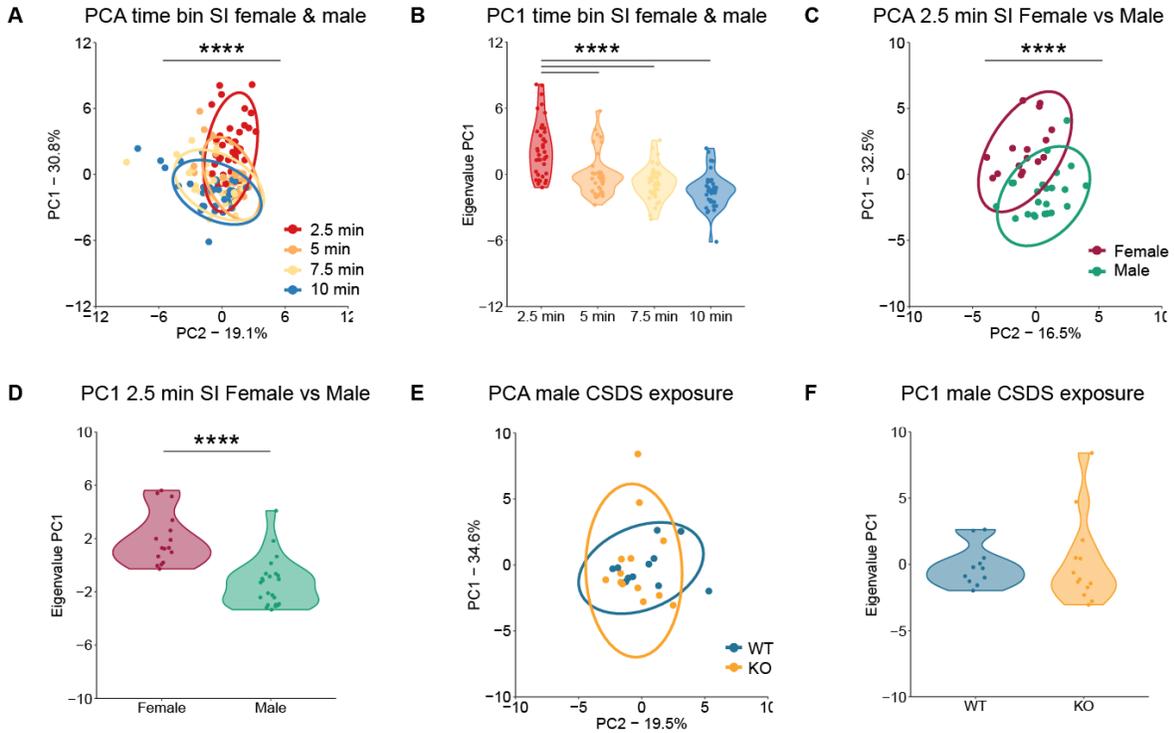
33. Oyarzabal, E. A. *et al.* Chemogenetic stimulation of tonic locus coeruleus activity strengthens the default mode network. *Sci Adv* **8**, (2022).
34. Zerbi, V. *et al.* Rapid Reconfiguration of the Functional Connectome after Chemogenetic Locus Coeruleus Activation. *Neuron* **103**, 702–718.e5 (2019).
35. McCall, J. G. *et al.* Locus coeruleus to basolateral amygdala noradrenergic projections promote anxiety-like behavior. *Elife* **6**, (2017).
36. Borodovitsyna, O., Flamini, M. D. & Chandler, D. J. Acute Stress Persistently Alters Locus Coeruleus Function and Anxiety-like Behavior in Adolescent Rats. *Neuroscience* **373**, 7–19 (2018).
37. Zitnik, G. A., Curtis, A. L., Wood, S. K., Arner, J. & Valentino, R. J. Adolescent Social Stress Produces an Enduring Activation of the Rat Locus Coeruleus and Alters its Coherence with the Prefrontal Cortex. *Neuropsychopharmacology* **41**, 1376–1385 (2016).
38. Finnell, J. E. *et al.* The contribution of the locus coeruleus-norepinephrine system in the emergence of defeat-induced inflammatory priming. *Brain Behav Immun* **79**, 102–113 (2019).
39. Chen, P. & Hong, W. Neural Circuit Mechanisms of Social Behavior. *Neuron* **98**, 16–30 (2018).
40. Sturman, O. *et al.* Deep learning-based behavioral analysis reaches human accuracy and is capable of outperforming commercial solutions. *Neuropsychopharmacology* **45**, 1942–1952 (2020).
41. Hånell, A. & Marklund, N. Structured evaluation of rodent behavioral tests used in drug discovery research. *Front Behav Neurosci* **8**, (2014).
42. Bordes, J., Miranda, L., Müller-Myhsok, B. & Schmidt, M. V. Advancing social behavioral neuroscience by integrating ethology and comparative psychology methods through machine learning. *Neurosci Biobehav Rev* **151**, 105243 (2023).
43. Lauer, J. *et al.* Multi-animal pose estimation, identification and tracking with DeepLabCut. *Nat Methods* **19**, 496–504 (2022).
44. Mathis, A. *et al.* DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat Neurosci* **21**, 1281–1289 (2018).
45. Bordes, J. *et al.* Automatically annotated motion tracking identifies a distinct social behavioral profile following chronic social defeat stress. *bioRxiv* 2022.06.23.497350 (2022) doi:10.1101/2022.06.23.497350.
46. Rodríguez, C. I. *et al.* High-efficiency deleter mice show that FLPe is an alternative to Cre-loxP. *Nat Genet* **25**, 139–140 (2000).
47. Häusl, A. S. *et al.* The co-chaperone Fkbp5 shapes the acute stress response in the paraventricular nucleus of the hypothalamus of male mice. *Mol Psychiatry* **26**, 3060–3076 (2021).
48. Gong, S. *et al.* Targeting Cre Recombinase to Specific Neuron Populations with Bacterial Artificial Chromosome Constructs. *The Journal of Neuroscience* **27**, 9817–9823 (2007).

49. Wagner, K. V *et al.* Differences in FKBP51 Regulation Following Chronic Social Defeat Stress Correlate with Individual Stress Sensitivity: Influence of Paroxetine Treatment. *Neuropsychopharmacology* **37**, 2797–2808 (2012).
50. Nagler, J., Schriever, S. C., De Angelis, M., Pfluger, P. T. & Schramm, K.-W. Comprehensive analysis of nine monoamines and metabolites in small amounts of peripheral murine (C57Bl/6 J) tissues. *Biomedical Chromatography* **32**, e4151 (2018).
51. Anderzhanova, E. *et al.* The stress susceptibility factor FKBP51 controls S-ketamine-evoked release of mBDNF in the prefrontal cortex of mice. *Neurobiol Stress* **13**, 100239 (2020).
52. Bonapersona, V. *et al.* The mouse brain after foot shock in four dimensions: Temporal dynamics at a single-cell resolution. *Proceedings of the National Academy of Sciences* **119**, (2022).
53. Martinez, M., Calvo-Torrent, A. & Herbert, J. Mapping Brain Response to Social Stress in Rodents With c-fos Expression: A Review. *Stress* **5**, 3–13 (2002).
54. Cullinan, W. E., Herman, J. P., Battaglia, D. F., Akil, H. & Watson, S. J. Pattern and time course of immediate early gene expression in rat brain following acute stress. *Neuroscience* **64**, 477–505 (1995).
55. Kovács, K. J. c-Fos as a transcription factor: a stressful (re)view from a functional map. *Neurochem Int* **33**, 287–297 (1998).
56. Reyes, B. A. S. *et al.* Neurochemically distinct circuitry regulates locus coeruleus activity during female social stress depending on coping style. *Brain Struct Funct* **224**, 1429–1446 (2019).
57. Nold, V. *et al.* Impact of Fkbp5 × early life adversity × sex in humanised mice on multidimensional stress responses and circadian rhythmicity. *Mol Psychiatry* **27**, 3544–3555 (2022).
58. Barsegyan, A., McGaugh, J. L. & Roozendaal, B. Noradrenergic activation of the basolateral amygdala modulates the consolidation of object-in-context recognition memory. *Front Behav Neurosci* **8**, (2014).
59. Belau, D. & McGaugh, J. Enhancement of extinction memory consolidation: The role of the noradrenergic and GABAergic systems within the basolateral amygdala. *Neurobiol Learn Mem* **86**, 123–132 (2006).
60. LaLumiere, R. T., Buen, T.-V. & McGaugh, J. L. Post-Training Intra-Basolateral Amygdala Infusions of Norepinephrine Enhance Consolidation of Memory for Contextual Fear Conditioning. *The Journal of Neuroscience* **23**, 6754–6758 (2003).

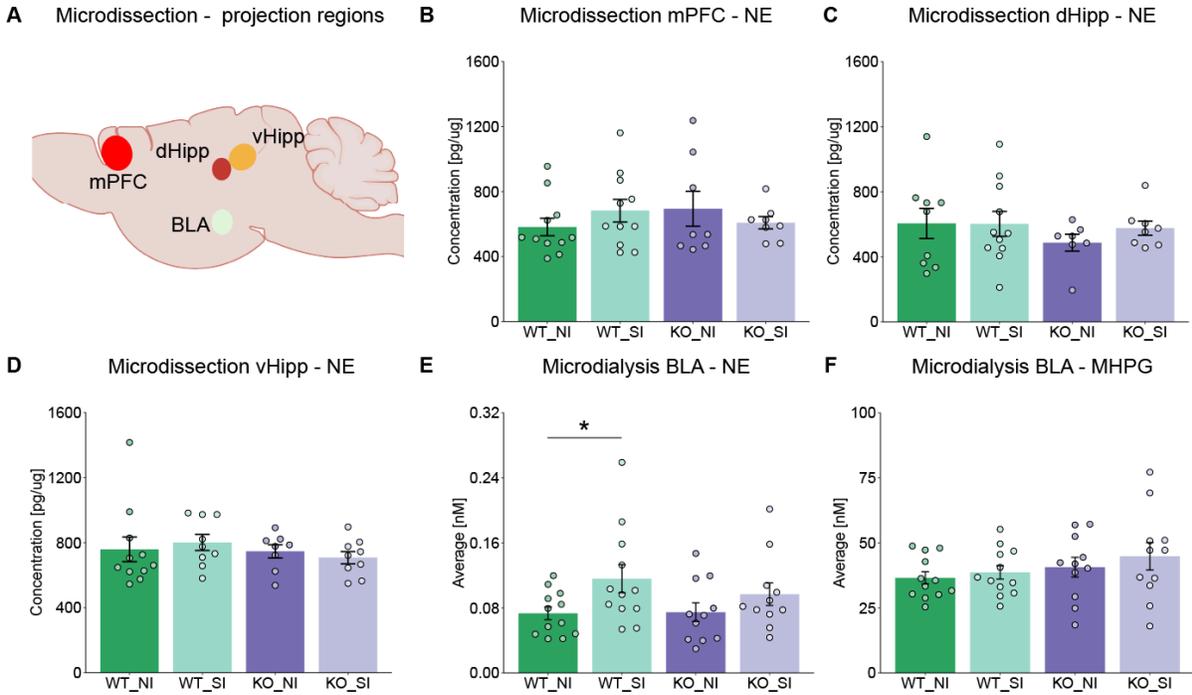
Supplemental materials



Supplemental figure 1. *Fkbp5* expression in the Dorsal raphe nucleus. A) *Fkbp5* mRNA expression in the DRN showed a significant main effect with the one-way ANOVA: $F(5,42)=12.85$, $p=1.33 \times 10^{-7}$. Post-hoc BH revealed that *Fkbp5* mRNA expression in the DRN was significantly increased in NS-ASDS compared to NS-NS ($p=8.35 \times 10^{-6}$) and NS-Res ($p=8.35 \times 10^{-6}$). The CSDS-ASDS mice also did have significantly increased *Fkbp5* levels compared to CSDS-NS ($p=0.00017$), or CSDS-Res ($p=0.0016$). The NS-ASDS mice did not have altered *Fkbp5* levels compared to CSDS-ASDS ($p=0.34$). **B)** Example in-situ hybridization scans in the DRN. In panel A; $n=8$ for all groups.



Supplemental figure 2. PCA analysis for the female and male social interaction task. **A)** The PCA time bin analysis comparing 4 bins revealed that the first 2.5 min time bin is significantly different from the other time bins. (Kruskal-Wallis test: $H(3)=56.00$, $p=4.21 \times 10^{-12}$). **B)** The PC1 eigenvalues of the SI time bin PCA. Post-hoc Wilcoxon: 2.5 min vs. 5 min ($W=1191$, $p=8.72 \times 10^{-6}$), 2.5 min vs. 7.5 min ($W=1321$, $p=1.91 \times 10^{-9}$), 2.5 min vs. 10 min ($W=1415$, $p=2.33 \times 10^{-13}$). **C-D)** The PCA on the 2.5 min SI data revealed a significant difference between sexes (Two-tailed Wilcoxon test: $W=339$, $p=1.17 \times 10^{-6}$). **E-F)** The PCA on the 2.5 min SI data for males showed no differences between WT and KO after CSDS exposure (Two-tailed Wilcoxon test: $W=82$, $p=0.57$). In panels A-B; $n=39$ across all four time bins. In panels C-D; $n=16$ for female and $n=23$ for male. In panels E-F; $n=11$ for WT and $n=13$ for KO.



Supplemental figure 3. Noradrenergic signaling between different project regions. **A)** Schematic overview of analyzed project regions of the LC with microdissection **B)** Microdissection measuring total NE levels in the mPFC. No main effect differences were observed using the two-way ANOVA: social interaction: $F(1,34)=0.1$, $p=0.75$, genotype: $F(1,34)=0.07$, $p=0.79$, social interaction*genotype: $F(1,34)=1.73$, $p=0.20$. **C)** Microdissection measuring total NE levels in the dHipp. No main effect differences were observed using the two-way ANOVA: social interaction: $F(1,31)=0.27$, $p=0.61$, genotype: $F(1,31)=0.86$, $p=0.36$, social interaction*genotype: $F(1,31)=0.38$, $p=0.54$. **D)** Microdissection measuring total NE levels in the vHipp. No main effect differences were observed using the two-way ANOVA: social interaction: $F(1,33)=0.0001$, $p=0.99$, genotype: $F(1,33)=0.81$, $p=0.37$, social interaction*genotype: $F(1,33)=0.50$, $p=0.49$. **E)** Microdialysis measuring NE release in the BLA. Posthoc BH revealed that WT social interaction (WT_SI) has significantly increased NE concentrations in the BLA compared to WT no interaction (WT_NI), $F(1,22)=4.99$, $p=0.036$, which was not altered in the *Fbcp5^{Nat}* background, comparing KO_NI with KO_SI, $p=0.23$. Two-way ANOVA: social interaction; $F(1,42)=6.22$, $p=0.017$, genotype; $F(1,42)=0.46$, $p=0.50$, social interaction*genotype; $F(1,42)=0.59$, $p=0.45$. **F)** Microdialysis measuring MHPG release in the BLA. No main effects were observed using the two-way ANOVA: social interaction; $F(1,42)=0.75$, $p=0.39$, genotype; $F(1,42)=2.03$, $p=0.16$, social interaction*genotype; $F(1,42)=0.085$, $p=0.77$. In panel B; $n=11$ for WT_NI, $n=11$ for WT_SI, $n=8$ for KO_NI, and $n=8$ for KO_SI. In panel C; $n=9$ for WT_NI, $n=11$ for WT_SI, $n=7$ for KO_NI, and $n=8$ for KO_SI. In panel D; $n=11$ for WT_NI, $n=9$ for WT_SI, $n=8$ for KO_NI, and $n=9$ for KO_SI. In panels E-F; $n=12$ for WT and $n=11$ for KO.

Chapter 7

Increasing resolution in stress neurobiology: from single cells to complex group behaviors

Miranda, L.*, **Bordes, J.***, Gasperoni, S., Lopez, J.P.

*Shared authorship

Originally published in:

Stress, 2023, 26(1), 2186141

Increasing resolution in stress neurobiology: from single cells to complex group behaviors

Lucas Miranda, Joeri Bordes, Serena Gasperoni & Juan Pablo Lopez

To cite this article: Lucas Miranda, Joeri Bordes, Serena Gasperoni & Juan Pablo Lopez (2023) Increasing resolution in stress neurobiology: from single cells to complex group behaviors, *Stress*, 26:1, 2186141, DOI: [10.1080/10253890.2023.2186141](https://doi.org/10.1080/10253890.2023.2186141)

To link to this article: <https://doi.org/10.1080/10253890.2023.2186141>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 09 Mar 2023.



Submit your article to this journal [↗](#)



Article views: 198



View related articles [↗](#)



View Crossmark data [↗](#)

Increasing resolution in stress neurobiology: from single cells to complex group behaviors

Lucas Miranda^{a,b,*} , Joeri Bordes^{c,*} , Serena Gasperoni^d  and Juan Pablo Lopez^d 

^aDepartment of Statistical Genetics, Max Planck Institute of Psychiatry, Munich, Germany; ^bInternational Max Planck Research School for Translational Psychiatry (IMPRS-TP), Munich, Germany; ^cResearch Group Neurobiology of Stress Resilience, Max Planck Institute of Psychiatry, Munich, Germany; ^dDepartment of Neuroscience, Karolinska Institutet, Stockholm, Sweden

ABSTRACT

Stress can have severe psychological and physiological consequences. Thus, inappropriate regulation of the stress response is linked to the etiology of mood and anxiety disorders. The generation and implementation of preclinical animal models represent valuable tools to explore and characterize the mechanisms underlying the pathophysiology of stress-related psychiatric disorders and the development of novel pharmacological strategies. In this commentary, we discuss the strengths and limitations of state-of-the-art molecular and computational advances employed in stress neurobiology research, with a focus on the ever-increasing spatiotemporal resolution in cell biology and behavioral science. Finally, we share our perspective on future directions in the fields of preclinical and human stress research.

ARTICLE HISTORY

Received 29 November 2022
Accepted 23 February 2023

KEYWORDS

Stress neurobiology; animal models of stress; single cells; complex behaviours; molecular neurobiology; computational science

Introduction

Stress is an important risk factor in the development of neuropsychiatric disorders including major depression, anxiety, post-traumatic stress disorder (PTSD), and other mood disorders (Davis et al., 2017; Musazzi et al., 2018; Musazzi & Marrocco, 2016; Sanacora et al., 2022). Elucidating the underlying cellular and molecular mechanisms responsible for the pathophysiology of psychiatric disorders requires the generation and implementation of preclinical animal models. Although unable to fully recapitulate the multidimensionality and complexity of stress-related psychiatric disorders in humans, they represent valuable tools to shed light onto the mechanisms underlying mental health disorders and develop appropriate pharmacological strategies.

Unraveling the complexity of the neurobiological circuits and molecular pathways underlying a healthy or abnormal stress response requires the combination and integration of cellular, molecular, and behavioral data. While traditional approaches lack in-depth spatial and temporal resolution, recent technological advancements have made it possible to improve these aspects considerably (Gururajan et al., 2018). For instance, single-cell transcriptomics allows to probe thousands of genes simultaneously and to dissect the contribution of distinct cell types involved in the stress response. Likewise, the implementation of activity-dependent labeling methods combined with brain clearing techniques, enables to ascertain which cells are activated following specific stressors, and to reconstruct the brain circuits involved in a

specific stress-response. Like all “omics” and high-throughput techniques, the implementation of these strategies generates large amounts of data. It is thus fundamental that they are complemented by appropriate computational and statistical tools. As a consequence, the advancement in molecular and cellular neuroscience techniques prompted a growth in the fields of computational science and the development of suitable data analysis software (Wang et al., 2020). In turn, the remarkable computational innovation stimulated a paradigm shift in the context of behavioral phenotyping, bringing about methods to automatically detect and analyze behaviors of interest (Shemesh et al., 2013). This now makes it possible to assess at the behavioral level the specific effect of different types of stressors (e.g. physical, psychological), stress paradigms (acute, chronic), developmental ages (e.g. early life, adolescence, adulthood, old age), and sex (males and females) in a time-effective manner, while considerably reducing manual scoring-related bias. In this commentary, we explore strengths and limitations of state-of-the-art methodologies employed in the field of stress neurobiology, focusing on molecular (*in vivo*) techniques, as well as computational (*in silico*) tools for both single-cell transcriptomic data analysis and automatic behavioral tracking systems, with an emphasis on the ever-increasing spatiotemporal resolution (Figure 1). While far from devoid of problems, we believe that the correct integration of molecular and computational techniques will greatly contribute to elucidating the role of stress in

CONTACT Juan Pablo Lopez  jpablo.lopez@ki.se  Karolinska Institutet, Solnavägen 9, Stockholm, 171 65, Sweden

*These authors contributed equally to this work

Co-first authorship order was determined by who got the most golden team shields on a set of 15 Panini FIFA World Cup 2022 sticker packs

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

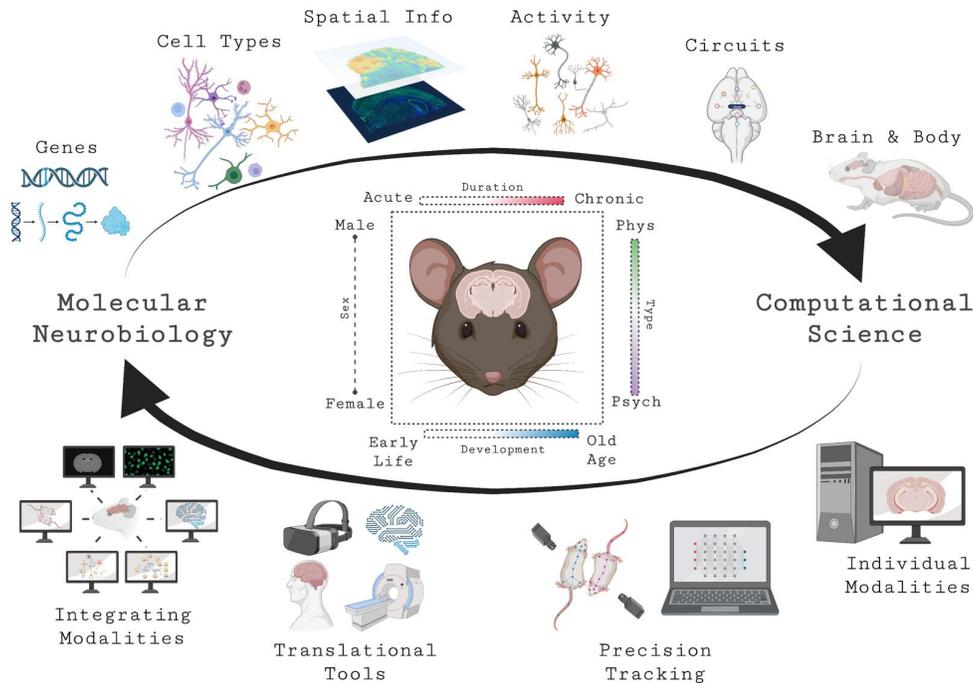


Figure 1. From macro to micro: increasing resolution in stress neurobiology. State-of-the-art molecular and computational advances employed in stress neurobiology research, with a focus on the ever-increasing spatiotemporal resolution in cell biology and behavioral science. The response to stress can be explored at different levels. For example, the type of stressor (physical [Phys] vs. psychological [Psych]), duration of the stressor (acute vs. chronic), developmental stage (early life, adolescence, adulthood, old age), or sex (male vs. female). Created with [BioRender.com](https://www.biorender.com).

neuropsychiatric disorders and in designing suitable treatment options.

Molecular neurobiology

Single cells: increasing spatial and temporal resolution of the brain

Cells are the essential building blocks of life and are therefore a crucial component to understand the biological mechanisms responsible for health and disease. Understanding the molecular profile of individual cells, for instance on the RNA or protein level, will enhance our understanding of the mechanisms by which stressors are perceived and processed into molecular, neuroendocrine, and behavioral responses under healthy and pathological conditions (Gururajan et al., 2018). In this section, we discuss the current state of the field of stress neurobiology from a molecular neurobiology perspective and provide our view, as early career scientists, on future directions.

First, we examine the ever-increasing resolution of the field at the micro level, in which there is an increased emphasis on information about genes, their cellular entity, and their surroundings. Currently, there is a wide variety of different techniques to investigate the molecular profile of cells. Traditional approaches include western blot, immunohistochemistry, *in situ* hybridization, and quantitative real-time PCR, among others. Over the years, these techniques have provided great insights into many cellular processes associated with a response to stress across different organs and brain regions. However, these techniques are limited to a small number of genes or proteins that can be measured

by a single experiment and require large amounts of input material. Over the last 20 years, the development of RNA sequencing technologies allowed for the quantification of thousands of genes in a single experiment by genome-wide analyses, shaping our understanding of the intricate mechanisms of the stress response in human and animal tissues. RNA sequencing has triggered a paradigm shift, in which a hypothesis-generating approach is utilized to investigate the role of novel molecular targets and their link to stress-related disorders. In addition, RNA sequencing techniques allow for high-throughput quantification of a variety of RNA species, including long-noncoding RNAs and microRNAs, which have now emerged as important regulators of the physiological response to a stressor (Issler et al., 2020; Lin et al., 2021; Lopez et al., 2018). However, these experiments require large amounts of RNA, from thousands of cells, and unfortunately, lack cell-type specificity. New developments in the field of molecular genomics now allow for single-cell and cell-type specificity using single-cell RNA sequencing (scRNA-seq) (Wen et al., 2022), which has emphasized and highlighted the contributions of different cell types in relation to a stressful event. For example, in 2021 we performed a cell type-specific, molecular characterization of all three components of the hypothalamic-pituitary-adrenal (HPA) axis, under baseline and chronic stress conditions (Lopez et al., 2021). In contrast to standard “bulk” RNA-sequencing methods, scRNA-seq allowed us to perform an unbiased characterization of cell types from the paraventricular nucleus of the hypothalamus (PVN), pituitary, and adrenal glands. We use the term “unbiased” here because currently, most cell-types are classified using a handful of established marker genes from the

literature, rather than using a comprehensive strategy that allows for an objective classification of different cell types that is based on unique transcriptional signatures from hundreds/thousands of genes. It is possible that a more reliable molecular signal of stress response remains elusive because of lack of an unbiased classification of cell types and the absence of better markers for their identification. Since then, other studies have demonstrated the importance of understanding how different cell types respond to an acute or chronic stressor (Dournes et al., 2020; Häusl et al., 2021; Kwon et al., 2022; Short, 2021; von Ziegler et al., 2022). Most importantly, these studies have provided extensive datasets as valuable resources for researchers and clinicians interested in the organism's nervous and endocrine responses to stress and the interplay between these tissues. Nevertheless, it is important to point out that scRNA-seq technologies are not without limitations. For example, the results from these experiments are often contingent on analysis parameters and other unaccounted variables, such as sample preparation, dissociation protocols, as well as proportions and sensitivities of cell types. Therefore, it is highly recommended to engage in independent validation of the primary findings to ensure that robust and consistent findings are reported. This may be cost prohibitive, but it is important as many investigators may use these datasets to generate new hypotheses and interpret previous findings. Another limitation of scRNA-seq is the depth of sequencing, which is much less than bulk RNA-seq. An alternative approach to explore cell-type specificity is to isolate a population of cells using a specific marker. This can be accomplished using flow cytometry or antibody-bound methods to enrich for a target population and sequence at greater depth, using bulk RNA-seq. However, it is important to point out that, this approach requires prior knowledge of target cell-type markers, and lacks the exploratory capabilities of single-cell experiments.

Unfortunately, due to the dissociation of single cells from target tissues, standard scRNA-seq techniques lack spatial information (Tian et al., 2022). Over the years, many studies in the stress field have shown the importance of cellular organization and function using well-established techniques, such as immunohistochemistry (IHC) (Hamilton et al., 2018) and fluorescent *in-situ* hybridization (FISH) (Engelhardt et al., 2021). However, these techniques only allow for a small number of genes to be detected and quantified and thus require specific hypothesis-driven gene targeting approaches. Interestingly, current developments in FISH have allowed for a significant increase in the number of targets, with more than 2,000 genes that can be labeled in one slice using enhanced electric FISH (Borm et al., 2022). Furthermore, recent improvements in sequencing methods now allow for single-cell transcriptomic analyses with spatial information and resolution (Moffitt et al., 2022). Importantly, these new technologies do not aim at replacing non-spatial techniques, but can often be seen as complementary. Along these lines, spatial transcriptomics in particular can also help annotate already available single-cell expression data. For example, Maynard et al. (2021) analyzed gene expression across the six layers of the human dorsolateral prefrontal cortex. They not only identified genes that were differentially expressed in

specific layers, but also used their data to improve the annotation of previously obtained, non-spatial datasets. Approaches like these could add information on existing data in other regions of the brain, the HPA axis, and the immune system, to name a few. Considering the advancements in FISH and scRNA-seq, in the future we can expect a significant increase in the molecular resolution at which we can assess how stress exposure influences changes in the expression of genes and their respective cell types.

A major drawback of these techniques is that the main outcome of the experiment remains a snapshot of the stress response in a tissue of interest, at a specific moment in time. This is a critical limitation, as the effects of stress exposure can vary substantially across different time points. Most importantly, these techniques cannot distinguish between those cells which are engaged directly during and after exposure to stress to those that remain unengaged. Obtaining activity-dependent information will be critical when investigating the response to a stressor. In the next section, we highlight several techniques that have been developed to capture the activity-dependent state of cells after exposure to a stimulus, within and across brain regions.

Activity sensors: understanding the individual role of brain cells and circuits in stress

Lack of spatiotemporal resolution of neuronal activity is a major problem for the precise dissection of brain circuits (Gururajan et al., 2018). Exposure to a stressful event triggers cellular activation in multiple temporal waves across different cells within a set of brain regions, which ultimately drives a neuroendocrine and behavioral response. The activity state of cells is an important proxy to investigate the cellular response system (Kawashima et al., 2014). Stressors can activate a spatially scattered subset of cells within homologous brain regions, which emphasizes the importance to distinguish cells based on their activity patterns. In response to cellular activation, different cell populations will use electrical and chemical synapses to communicate with other cells. Chemical synapses release one or several different neurotransmitters (NTs) and neuromodulators (NMs), many of which are related to the stress response system, such as norepinephrine and corticotropin-releasing factor (Deussing & Chen, 2018; Hökfelt et al., 2018). Several techniques have been developed to explore the activity of neuronal networks, such as microdialysis and mapping of brain networks using immediate early genes. These techniques have provided important insights into the different brain regions activated in response to specific stressful events, but have limited cell-type specificity and high spatiotemporal resolution.

Promising and recently developed techniques are now aiming to provide new information to explore the activity of individual cells and neuronal circuits within a network. For example, genetically encoded GPCR activation-based (GRAB) sensors (Feng et al., 2019), reviewed by Wu et al. (2022) to investigate *in-vivo* fluctuations of neurotransmitters (NTs) and neuromodulators (NMs). GRAB sensors are highly selective to the NT or NM of interest, and upon binding will

change their conformation, so a fluorescent signal can be detected. The GRAB sensors can target specific cell populations by using cre-dependent labeling, which opens up the possibility to investigate the contributions of different cell types within the system. In addition, they are able to detect NT and NM fluctuations within the millisecond time window. This makes the GRAB sensors a strong tool for investigating the cellular response to stress in the brain, as it has high molecular selectivity and temporal sensitivity. Another promising technique is the implementation of activity-dependent labeling methods. Genetic labeling of neurons, with a specific response feature, is an emerging technology for the precise dissection of functionally heterogeneous brain circuits. Immediate early gene mapping has been widely used for decades to identify brain regions that are activated by external stimuli, however high spatiotemporal resolution has proved to be time consuming and extremely laborious (Franceschini et al., 2020). A recent study used multiple cohorts of mice, sacrificed at different time points, after exposure to a particular stressor and highlighted the importance of timing (the temporal component) in stress research, as they observed a specific time-dependent pattern of c-Fos protein expression across different brain regions (Bonapersona et al., 2022). However, the expression of c-Fos is ubiquitous across neuronal populations (Cruz-Mendoza et al., 2022), which limits the information that can be gathered regarding specialized functions of particular neuronal types. In addition, the statistical analysis for such brain-wide analyses using different time-dependent cohorts is complex and highly variable due to the individual differences between cohorts. The recent characterization of the promoter and enhancer elements responsible for neuronal activity-dependent transcription has opened new avenues for the dissection of active neurons, allowing for the characterization of neural ensembles and circuits in greater detail (Kawashima et al., 2014). Using activity-driven labeling, it is now possible to label and track activated cell populations in a specific time window through the brain using viruses and genetic mouse lines, such as the enhanced synaptic activity responsive element E-SARE (Kawashima et al., 2013) or targeted recombination in active populations (TRAP) and TRAP2 (Allen et al., 2017). Most importantly, these innovative techniques can be combined with *in-vivo* tracking tools, such as electrophysiology, optogenetics, DREADDs, calcium imaging, as well as GRAB sensors to provide a deeper understanding of a healthy and abnormal stress response. Currently, only a handful of studies have investigated the stress response system using activity-dependent labeling and *in-vivo* tracking tools (Koutlas et al., 2022; Niu et al., 2022; Ramirez et al., 2015). For example, an interesting study using first-generation TRAP mice was conducted by Ramirez and colleagues, who showed that the reactivation of dentate gyrus cells, which were previously labeled during a positive experience, can rescue stress-induced depression-like behaviors (Ramirez et al., 2015). However, more recent studies utilize a new generation of TRAP mice (TRAP2), which allow for a more specific signal, only in neurons of interest. More specifically, using TRAP2 mice, Koutlas and colleagues showed that stress-activated neurons in the ventral tegmental area have different

electrophysiological properties, as compared to non-activated neurons in the same region (Koutlas et al., 2022). Furthermore, combining calcium imaging tools with activity-dependent labeling would allow for exclusive investigation of cellular plasticity from stress-responsive cells and exploration of their activity properties at different time scales, from immediate (acute) to long-term (chronic) effects. An important remark is that the “tagging” of different cell types simultaneously is not possible, which limits the identification and contributions of cell type-specific effects.

While current advances in the field of neurobiology using FISH, scRNA-seq, GRAB sensors, and activity-dependent labeling methods have been aimed at increasing molecular resolution (the micro level), these techniques by themselves do not inform at the level of circuits and networks, as well as interactions across brain regions and communication with other peripheral systems (the macro level). These topics will be discussed in the next section.

Brain and body: investigating whole systems to better understand the stress response

A stressful event triggers a cascade of cellular responses in many different brain regions, and peripheral systems, which in turn influence each other (Dedic et al., 2018). It is crucial to consider the entire brain and body as a holistic entity, to further understand different systems and characterize novel pathways related to stress exposure and response. Biochemical and neuroendocrine data, such as circulating levels of glucocorticoids (GCs) have been used as an important parameter to measure the stress response in animal models and humans. Nowadays, advances in multiplex immunoassays can provide a more holistic view of biological markers (e.g. GCs, cytokines, catecholamines, vasopressin, among others), and even distinguish markers related to different types of stressors (e.g. acute versus chronic stress) (Ataollahi et al., 2022, Tighe et al., 2015). In addition, several physiological measurements are now used to investigate how stress responses can promote energy reallocation to support survival. For example, metabolic cages allow for quantification and exploration of several physiological parameters, such as weight, respiratory exchange rate, and energy expenditure, which have been found to be differentially altered between different stress paradigms (Kuti et al., 2022). Most of these measurements are readily available from numerous human psychiatric and preclinical studies. Being able to integrate this data, into the spectrum of single cell – whole brain studies can increase translatability across species and studies.

Moreover, a possible way to obtain a more comprehensive view of stressed-induced alterations in the brain is by using a series of slices through the entire brain and labeling the expression of immediate-early genes using IHC or FISH, which has provided interesting insights into a brain-wide analysis of different cellular targets related to stress exposure (Scharf et al., 2011; Silva et al., 2019). However, as previously stated, these approaches are extremely laborious, time-consuming, and are limited to a small number of genes that can be detected and quantified. Another approach is the use of

magnetic resonance imaging (MRI), in which whole brain activity can be obtained in a single experimental setup, which provides insight into the activation and communication of particular regions across the brain (Mandino et al., 2019). Unfortunately, using MRI animals can only be tested under deep anesthesia, which severely limits external manipulations, such as natural exposure to stressors during experimental recordings. An alternative method to investigate whole-brain activity using cerebrovascular fluctuations is functional ultrasound imaging (fUS), which enables *in-vivo* recordings without anesthesia (Deffieux et al., 2021). Next to investigating the blood flow changes across brain regions, it is crucial to be able to investigate activity patterns across the brain at higher resolution in order to trace and investigate the activated circuits at the single-cell level, which cannot be achieved using techniques, such as MRI and fUS. A technique that can explore cellular activity across the entire brain, while providing single-cell resolution is brain clearing. Brain clearing has been rapidly advancing with different methods, such as CLARITY (Chung et al., 2013) and iDISCO (Renier et al., 2014). These techniques help us visualize protein expression throughout a cleared tissue, such as the entire brain, at an incredible cellular resolution. Nevertheless, we believe that to successfully capture the complexity of the stress response, the combination of these different techniques will be crucial. For instance, combining cell type specific methods, with activity-driven labeling tools, and brain clearing techniques will provide a more well-rounded view of the brain during or after exposure to stress. Ultimately, this will allow us to use a more unbiased method to investigate specific brain regions, cell types, and cell populations related to the stress response system. An excellent showcase for combining these tools are recent studies published by Niu *et al.*, in 2022 (Niu et al., 2022), and Davoudian et al., 2023 (Davoudian et al., 2023). In the first, the authors start their study using whole-brain imaging after restraint stress then narrowed their focus to a few identified stress-responsive regions, including a novel target in the claustrum. Subsequently, they labeled a stress-responsive neuronal ensemble in the claustrum, using activity-dependent labeling tools and observed that the silencing of this neuronal network, using DREADDS, resulted in attenuation of anxiety-related behaviors, whereas the activation of the same network elicited those behaviors. Similarly, in 2023 Davoudian and colleagues employed whole-brain serial two-photon microscopy and light sheet microscopy to map the expression of the immediate early gene, *c-Fos*, in male and female mice, following administration of ketamine and psilocybin. Their systematic mapping approach produced an unbiased list of brain regions impacted by both treatments.

Furthermore, in the future important topics, such as the influences of sex on the stress response system, the molecular mechanisms and circuits involved in treatment response, or the connection between the central and peripheral nervous systems can be investigated in greater detail using such an approach. For reference, Brivio *et al.*, summarize most of the studies that have established sex differences in the neurobiological and behavioral effects of stress exposure (Brivio et al., 2020). These new tools will significantly improve our understanding of the molecular mechanisms and cellular

circuits responsible for the development of stress-related psychiatric disorders. However, the generation and analysis of these increasingly more complex and larger datasets have created great statistical and computational challenges in our field, hence the need for the development and integration of computational tools in the field of stress neurobiology.

Computational science

Digging deeper: leveraging computational advances to increase resolution in individual data modalities

Many of the questions that the stress neurobiology field is currently trying to address require a joint collection of many data modalities to reach sound conclusions. As technology advances, more data becomes available in different areas such as genomics, transcriptomics, proteomics, circuits, and behavior, to name a few. This renders an apparent need for developing standardized ways of taking advantage of this increased resolution, without losing sight of the big picture (*their interaction*). Furthermore, not only does this increase in resolution and data volume have value in itself, but it also carries the potential to incentivize the development of new tools that leverage computational developments, tailored to the tasks at hand. For example, the field of behavioral and computational science has witnessed an increasing number of statistical and machine learning tools designed to tackle different arising problems and automate laborious tasks, which has a huge impact in how research is being done to study the molecular mechanisms and behaviors associated with a stress response. In this section, we discuss the current state of the computational field, from a stress research perspective, and illustrate our view on where we think research could move next.

We will start by discussing the field at the micro level (that is, increased resolution in molecular data modalities) which allows us to inspect closer aspects of biology that were inaccessible before. As discussed earlier, increasing data volume and resolution in transcriptomics, has sparked a plethora of tools and methods that can make such analyses manageable. On the single-cell side, programs like SCANPY (Wolf et al., 2018) and SEURAT (Satija et al., 2015) have succeeded in making state-of-the-art processing and analysis accessible to a broad spectrum of researchers. To date, these have accumulated thousands of citations, and the user basis continues to grow. In addition, new extensions that handle new data modalities continue to be released and maintained, such as spatial transcriptomics (Palla et al., 2022), which is helping make unprecedented progress in the study of both tissue organization and cellular communication. Moreover, tools that leverage the ever-growing public datasets to automate even further workloads (for example, automatic cell annotation) are an example of the positive feedback loop these tools generate (Fischer et al., 2021). This level of standardization portrays substantial benefits for several related fields, and stress research is not an exception, with high implications for basic understanding of cell composition and gene expression in relevant tissues to novel drug targets and the development of new treatments. As an example of the

latter, in 2022, using a combination of automatic behavioral tracking techniques and state-of-the-art scRNA-seq methods, we identified cell-type-specific molecular signatures, and a previously unknown mechanism of action, for the sustained antidepressant effects of ketamine in glutamatergic neurons of the ventral hippocampus of adult mice (Lopez et al., 2022). We expect that, in the near future, these technologies will continue to shed light not only into cellular mechanisms underlying the action of drugs used in stress-related disorders, but also hint at new potential pharmacological targets that could be exploited in the future. Finally, while still unexplored in the stress field, to the best of our knowledge, the technical advances in spatial transcriptomics could accelerate these findings by providing access to crucial information on cellular distribution and communication within a given tissue.

So far, we have focused on areas in which breakthroughs in the experimental domain have triggered an increase in data volume, which in turn sparked the need for new computational approaches (either completely novel or borrowed from other computational and statistical fields). The case of behavioral analysis, however, follows the opposite trend: here, it was the thoughtful application of recent computational techniques, such as convolutional neural networks (CNNs) and other computer vision advances, which led to a rapid increase in data collection, and ultimately to a drastic change in how research is being carried out and the types of questions people can ask. In the next section, we discuss how precision behavioral tracking is an emerging and exacting new field in neuroscience research.

Precision tracking: automated systems to dissect the behavioral language of rodents in stress research

In 2013, Shemesh et al. (2013) developed an automatic phenotyping system based on video color recognition, known as the “Social Box” (SB). Here, the authors described how social behavior in mice develops in a semi-natural environment, using a set of techniques that aim to quantify behavioral traits in an automated way, thus freeing researchers from the burden of laborious manual quantification. The authors automatically tracked several groups of mice in their home environment and investigated how individual behavior is strongly interdependent in their groups. In a follow-up study in 2019, Forkosh and colleagues developed a model, using the SB system, that captures and outlines stable personality traits in mice (Forkosh et al., 2019). Although undoubtedly insightful, this work and many that followed (Anpilov et al., 2020; Forkosh et al., 2019; Karamihalev et al., 2020; Shemesh et al., 2016) were limited to tracking the central position of each animal. Furthermore, in this and other contemporary approaches, animal identification relied on dedicated (often expensive or invasive) hardware, such as radio frequency identification (RFID) or color hair dyes (Shemesh et al., 2013, de Chaumont et al., 2012).

Many of these issues were addressed in recent years by the development of neural network models that work on image data directly, without the need for physical markers. For example, tools such as DeepLabCut (DLC) (Mathis et al.,

2018), Social Leap Estimates Animal Poses (SLEAP) (Pereira et al., 2022), and SIPEC (Marks et al., 2022), have made it possible to gather enormous amounts of time series data on multiple body parts with human-level accuracy (Sturman et al., 2020). In addition, some of these models are now capable of retaining individual identification in social settings, without the need for dedicated hardware (Lauer et al., 2022). A concept we believe is worth mentioning here is *transfer learning*. That is, leveraging of previously trained models to classify gigantic datasets of unrelated images, which can lead to very good tracking with little (or no) labeling (known as few-shot learning) (Lauer et al., 2022, Ye et al., 2022). Furthermore, this has been shown to work well both in lab environments as in the wild, enabling its use for ethological studies. While marker-less animal tracking is in itself an accomplishment worth mentioning, many tools have become available that can take this one step further, and identify behavioral patterns in motion tracking data in both supervised (Nilsson, 2022; Segalin et al., 2021) and unsupervised (Bordes & Miranda, 2022; Hsu & Yttri, 2021; Luxem et al., 2022) ways, paving the way for automated behavioral screenings, that are both less laborious and more robust than more classical methods. Along these lines, we recently developed and introduced an open-source tool called DeepOF (Bordes & Miranda, 2022), which is capable of reporting interpretable patterns in open field motion tracking data in both supervised and unsupervised ways. The study emphasizes the importance of such analyses for stress research, as we showed how DeepOF can be used to detect distinct stress-induced behavioral patterns following chronic social defeat stress. In particular, we see (in a fully unsupervised way) an increase in huddling and escaping behaviors in chronically stressed mice, and an enrichment in exploratory patterns in controls. Moreover, DeepOF can detect habituation to non-hostile environments, reporting how behavioral differences between stressed and non-stressed mice fade over time. In a recent publication, Shemesh and Chen review different novel systems fit to investigate the behavior of rodents and discuss what they deemed as a paradigm shift in translational psychiatry through rodent neuroethology (Shemesh & Chen, 2023). The authors suggest that these new methods possess the best out of classical ethology and the reductive behaviorist approach, which may provide a breakthrough in discovering new efficient therapies for mental illnesses.

All these developments can have large implications for stress research. First, by measuring the position over time of one or more markers in less restricted environments, scientists can increase throughput, since extracting information from freely moving animals makes it easier to replace expensive and time-consuming batteries of univariate tests, while significantly reducing the large numbers of animals needed to accomplish the task. In addition, stress research is a field in which the leap between human conditions and their animal equivalents is often significant and questioned. Since these tools allow for more complex data-driven definitions of the outcomes we intend to measure, they have the potential to increase construct and face validity. For example, earlier this year we, as part of a larger group, proposed an algorithm to identify a ‘depression-like syndrome’ in mice based

on mappings from both DSM-V and ICD-11 (von Mücke-Heim et al., 2022). While rodent behavior will still remain a distant proxy of their human counterpart, given that factors such as social, economic, and inferential features are hard (if not impossible) to model, we believe efforts like this, which yield clearer, standardized definitions of preclinical phenotypes, will be extremely important for the field moving forward.

Translational tools: novel methods to improve translatability in stress research

All in all, both omics and motion tracking examples illustrate well how, in our view, having more data can lead to increased resolution and, in turn, accelerated discovery. However, to date, they are mostly applicable to animal models. Given that the focus of stress research is, at the end of the day, understanding and improving the life quality of people, translation and research in humans are of course crucial. In this regard, developments in understanding human behavior using virtual reality (VR) are worth mentioning.

VR currently allows researchers to track movement with precision in carefully created environments, making it possible to translate paradigms such as fear conditioning to human subjects in a noninvasive way (Binder et al., 2022, Binder & Spoomaker, 2020). Furthermore, imaging techniques like structural and functional magnetic resonance imaging (MRI), promise to accelerate translation by enabling data generation directly from human brains. While low test-retest reliability and potential construct validity issues coming from the heterogeneity of the labels that researchers use to study stress-related psychopathologies (Miranda et al., 2021), we believe that data-driven initiatives such as the research domain criteria (RDoC) (Insel et al., 2010, Morris et al., 2022), together with large scale multi-site data collection efforts (such as PRONIA) have enormous potential on bringing these promises closer to clinical reality (Haidl et al., 2023; Luutonen et al., 2013; Popovicet al., 2020). Furthermore, specific tools such as Neurominer (Koutsouleris, 2022), provide state of the art machine learning tools for brain imaging data with little-to-no code, which can be helpful in bringing this kind of expertise closer to doctors, in search for multivariate patterns that may aid diagnosis, prognosis prediction, and treatment optimization of stress-related disorders. Finally, while these developments have led on their own right to exciting research in the field of stress, they are limited to extending single data modalities. Understanding the stress response system goes far beyond understanding single cells, neural activity or behavior alone, and we believe that the key in the near future will rely on data modality integration.

Reaching broader: gaining integrated knowledge by combining multimodal data

A living system is far more than the sum of its parts, with different biological levels interacting and regulating one another constantly in complex ways. From genetics, transcriptomics, epigenomics, and proteomics, to neural signaling, behavior, and environmental factors, being able to merge

information acquired at different biological levels in clever ways can be key to understanding any phenotype (Stahlschmidt et al., 2022). This can help exploit the available data more efficiently, and lead to more holistic research questions. Moreover, a healthy response to stress depends on the interplay of many regulatory factors acting at several interdependent levels, which result in the allocation of energy resources to resolve the stressor situation (Russell & Lightman, 2019). Efforts in both describing this response, and understanding how it's altered in stress-related disorders in a multimodal way can help disentangling individual differences between, for example, susceptibility and resilience toward stress exposure or response and non-response to antidepressant treatments.

At a basic level, multimodal integration requires researchers to draw conclusions of experiments describing multiple (complementary) axes of the same problem, and drawing conclusions explaining all observed patterns. von Ziegler et al. (2022), for example, used a combination of proteomics, phospho-proteomics, bulk and single-nucleus RNAseq, and TRAP sequencing, to describe the temporal response in the mouse hippocampus to acute stress induced by forced swimming. By exploring all data types independently and taking prior knowledge into account, they provided a comprehensive analysis of the temporal dynamics involved. While undoubtedly useful, this approach may not scale to larger and more complex datasets, as researchers would have to learn their joint properties by hand. Furthermore, experiments may have different sensitivities, time scales, and intrinsic artifactual limitations, which in turn highlights the need for technologies capable of storing, handling, and automatically reporting joint features from multimodal data. Along these lines, several extremely relevant subfields for stress research, such as omics, are flourishing with options for researchers to benefit from. The recently published MUON package (Bredikhin et al., 2022), for example, aims at providing accessible and scalable storage and manipulation of multiple omics layers, where different modalities can be organized and analyzed with ease. The also recent tool MEFISTO (Velten et al., 2022), for instance, can then be used to map all modalities to a shared embedding space, using latent factor analysis. Interestingly, these tools are even capable of leveraging spatial and temporal dimensions, when available.

Another key point where advances are promising is the integration of behavioral and neural data. This remains key for studies going from basic neuroscience to psychiatric research (including stress), as finding neural correlates of adaptive behavioral patterns can pave the way to gain mechanistic insights into the mechanisms causing pathology or drug action. Along these lines, the recently presented software CEBRA (Schneider et al., 2022) promises to be of great help. Using a representation-learning approach, the package is able to report non-linear neural correlates of behavior, directly enabling questions regarding how one affects the other in complex ways that may be difficult to detect without computer assistance.

Concluding remarks

Here, we have discussed the strengths and limitations of state-of-the-art molecular and computational advances employed in stress neurobiology research, with a focus on the ever-increasing spatiotemporal resolution. Overall, we expect these types of molecular techniques and computational tools to encompass more combinations of modalities as the field matures, and increasing high-quality data becomes available. We want to highlight that with many technological advancements making the integration of these datasets possible, especially those involving complex, black-box models, *explainability* and *interpretability* of results is key to avoid reporting non-generalizable results that may ultimately be prejudicial to the field. While, in some cases, research can arguably directly inspect results visually (as it is the case for motion tracking), anything involving making predictions that rely on biological data should be thoroughly tested, especially if the underlying dataset is small or too specific, to make sure that our models are not learning from noise, or undetected confounders. Fortunately, the research community is becoming more aware of this issue, and both tools and best practices guidelines (Goodwin et al., 2022; Luecken & Theis, 2019) are being published to sort it out. As it is already the case with artificial intelligence in healthcare as a whole, we expect this topic to be on the spotlight of stress research as available tools become more complex. Finally, while the real impact on stress research remains to be explored, we strongly believe that integrating multimodal and complementary datasets will shed light over patterns too complex for humans to interpret directly, but relevant to ultimately understand and treat such a complex phenomenon. From single cells to social behavior, the dream of jointly mapping stress response as a whole is closer than ever.

Acknowledgements

The authors thank Dr. Bertram Müller-Myhsok, Dr. Mathias V. Schmidt, Dr. Patrizia Romualdi and Dr. Alon Chen for their mentorship and unconditional support. Figure 1 was created with BioRender.com

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

L.M. is supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Ph.D. grant agreement No. 813533. J.B. is supported by the "Kids2Health" grant of the Federal Ministry of Education and Research [01GL1743C]. J.P.L. is supported by Starting Grants from The Strategic Research Area Neuroscience (StratNeuro), The Svenska Sällskapet För Medicinsk Forskning (SSMF) [No SG-22-0204-H-02], and Strategic Recruitment in Research Funds from Karolinska Institutet and the Department of Neuroscience..

Notes on contributors

Lucas Miranda is a PhD student in the Statistical Genetics group of the Max Planck Institute of Psychiatry, directed by Dr. Bertram Müller-Myhsok

(Munich, Germany). His research focuses mainly on machine learning analysis of time series data, with a strong focus on developing tools for behavioral segmentation.

Joeri Bordes is a PhD student in the Neurobiology of Stress Resilience laboratory of Dr. Mathias Schmidt at the Max Planck Institute of Psychiatry (Munich, Germany). His research aims to implement novel behavioral analysis tools to unravel the influence of stress on social behavior and fear memory.

Serena Gasperoni is a PhD student in the laboratory of Dr. Juan Pablo Lopez in the Department of Neuroscience at Karolinska Institutet (Stockholm, Sweden). Her research aims to identify the molecular mechanisms, cellular circuits and behaviors associated with an antidepressant response to psychedelic compounds.

Dr. Juan Pablo Lopez is an Assistant Professor in the Department of Neuroscience at Karolinska Institutet (Stockholm, Sweden). His laboratory combines advanced molecular, behavioural and computational tools to characterize the molecular mechanisms, cellular circuits, and behavioural "language" associated with stress-related psychiatric disorders and their treatments, using animal models.

ORCID

Lucas Miranda  <http://orcid.org/0000-0002-5484-6744>
 Joeri Bordes  <http://orcid.org/0000-0003-2909-2976>
 Serena Gasperoni  <http://orcid.org/0000-0003-3333-9396>
 Juan Pablo Lopez  <http://orcid.org/0000-0002-5812-4220>

References

- Allen, W. E., DeNardo, L. A., Chen, M. Z., Liu, C. D., Loh, K. M., Fenno, L. E., Ramakrishnan, C., Deisseroth, K., & Luo, L. (2017). Thirst-associated preoptic neurons encode an aversive motivational drive. *Science*, 357(6356), 1149–1155. <https://doi.org/10.1126/science.aan6747>
- Anpilov, S., Shemesh, Y., Eren, N., Harony-Nicolas, H., Benjamin, A., Dine, J., Oliveira, V. E. M., Forkosh, O., Karamihalev, S., Hüttel, R.-E., Feldman, N., Berger, R., Dagan, A., Chen, G., Neumann, I. D., Wagner, S., Yizhar, O., & Chen, A. (2020). Wireless optogenetic stimulation of oxytocin neurons in a semi-natural setup dynamically elevates both pro-social and agonistic behaviors. *Neuron*, 107(4), 644–655.e647. <https://doi.org/10.1016/j.neuron.2020.05.028>
- Ataollahi, M., Nejad, J. G., & Park, K. H. (2022). Selection of appropriate biomatrices for studies of chronic stress in animals: A review. *Journal of Animal Science and Technology*, 64(4), 621–639. <https://doi.org/10.5187/jast.2022.e38>
- Binder, F. P., & Spoormaker, V. I. (2020). Quantifying human avoidance behavior in immersive virtual reality. *Frontiers in Behavioral Neuroscience*, 14, 569899. <https://doi.org/10.3389/fnbeh.2020.569899>
- Binder, F. P., Pöhlchen, D., Zwanzger, P., & Spoormaker, V. I. (2022). Facing your fear in immersive virtual reality: Avoidance behavior in specific phobia. *Frontiers in Behavioral Neuroscience*, 16, 827673. <https://doi.org/10.3389/fnbeh.2022.827673>
- Bonapersona, V., Schuler, H., Damsteegt, R., Adolfs, Y., Pasterkamp, R. J., van den Heuvel, M. P., Joëls, M., & Sarabdjitsingh, R. A. (2022). The mouse brain after foot shock in four dimensions: Temporal dynamics at a single-cell resolution. *Proceedings of the National Academy of Sciences of the United States of America*, 119(8), e2114002119. <https://doi.org/10.1073/pnas.2114002119>
- Bordes, J., & Miranda, L. (2022). Automatically annotated motion tracking identifies a distinct social behavioral profile following chronic social defeat stress. *BioRxiv*. <https://doi.org/10.1101/2022.06.23.497350>
- Borm, L. E., Mossi Albiach, A., Mannens, C. C. A., Janusauskas, J., Özgün, C., Fernández-García, D., Hodge, R., Castillo, F., Hedin, C. R. H., Villablanca, E. J., et al. (2022). Scalable in situ single-cell profiling by electrophoretic capture of mrna using eel fish. *Nature Biotechnology*, 41(2), 222–231. <https://doi.org/10.1038/s41587-022-01455-3>

- Bredikhin, D., Kats, I., & Stegle, O. (2022). Muon: Multimodal omics analysis framework. *Genome Biology*, 23(1), 42. <https://doi.org/10.1186/s13059-021-02577-8>
- Brivio, E., Lopez, J. P., & Chen, A. (2020). Sex differences: Transcriptional signatures of stress exposure in male and female brains. *Genes, Brain and Behavior*, 19(3), e12643.
- Chung, K., Wallace, J., Kim, S. Y., Kalyanasundaram, S., Andalman, A. S., Davidson, T. J., Mirzabekov, J. J., Zalocusky, K. A., Mattis, J., Denisin, A. K., et al. (2013). Structural and molecular interrogation of intact biological systems. *Nature*, 497(7449), 332–337. <https://doi.org/10.1038/nature12107>
- Cruz-Mendoza, F., Jauregui-Huerta, F., Aguilar-Delgado, A., García-Estrada, J., & Luquin, S. (2022). Immediate early gene c-fos in the brain: Focus on glial cells. *Brain Sciences*, 12(6), 687. <https://doi.org/10.3390/brainsci12060687>
- Davis, M. T., Holmes, S. E., Pietrzak, R. H., & Esterlis, I. (2017). Neurobiology of chronic stress-related psychiatric disorders: Evidence from molecular imaging studies. *Chronic Stress*, 1, <https://doi.org/10.1177/247054701710916>
- Davoudian, P. A., Shao, L. X., & Kwan, A. C. (2023). Shared and distinct brain regions targeted for immediate early gene expression by ketamine and psilocybin. *ACS Chemical Neuroscience*, 14(3), 468–480. <https://doi.org/10.1021/acscchemneuro.2c00637>
- de Chaumont, F., Coura, R. D., Serreau, P., Cressant, A., Chabout, J., Granon, S., & Olivo-Marin, J. C. (2012). Computerized video analysis of social interactions in mice. *Nature Methods*, 9(4), 410–417. <https://doi.org/10.1038/nmeth.1924>
- Dedic, N., Chen, A., & Deussing, J. M. (2018). The crf family of neuropeptides and their receptors - mediators of the central stress response. *Curr Mol Pharmacol*, 11(1), 4–31.
- Deffieux, T., Demené, C., & Tanter, M. (2021). Functional ultrasound imaging: A new imaging modality for neuroscience. *Neuroscience*, 474, 110–121. <https://doi.org/10.1016/j.neuroscience.2021.03.005>
- Deussing, J. M., & Chen, A. (2018). The corticotropin-releasing factor family: Physiology of the stress response. *Physiological Reviews*, 98(4), 2225–2286. <https://doi.org/10.1152/physrev.00042.2017>
- Dourmes, C., Chen, A., & PrDourmes, V. O. (2020). Hypothalamic glucocorticoid receptor in crf neurons is essential for hpa axis habituation to repeated stressor. *BioRxiv*.
- Engelhardt, C., Tang, F., Elkhateib, R., Bordes, J., Brix, L. M., van Doeselaar, L., Häusel, A. S., Pöhlmann, M. L., Schraut, K., Yang, H., et al. (2021). Fkbp51 in the oval bed nucleus of the stria terminalis regulates anxiety-like behavior. *eNeuro*, 8(6), ENEURO.0425-21.2021. <https://doi.org/10.1523/ENEURO.0425-21.2021>
- Feng, J., Zhang, C., Lischinsky, J. E., Jing, M., Zhou, J., Wang, H., Zhang, Y., Dong, A., Wu, Z., Wu, H., Chen, W., Zhang, P., Zou, J., Hires, S. A., Zhu, J. J., Cui, G., Lin, D., Du, J., & Li, Y. (2019). A genetically encoded fluorescent sensor for rapid and specific in vivo detection of norepinephrine. *Neuron*, 102(4), 745–761.e748. <https://doi.org/10.1016/j.neuron.2019.02.037>
- Fischer, D. S., Dony, L., König, M., Moede, A., Zappia, L., Heumos, L., Tritschler, S., Holmberg, O., Aliee, H., & Theis, F. J. (2021). Sfaira accelerates data and model reuse in single cell genomics. *Genome Biology*, 22(1), 248. <https://doi.org/10.1186/s13059-021-02452-6>
- Forkosh, O., Karamihalev, S., Roeh, S., Alon, U., Anpilov, S., Touma, C., Nussbaumer, M., Flachskamm, C., Kaplick, P. M., Shemesh, Y., & Chen, A. (2019). Identity domains capture individual differences from across the behavioral repertoire. *Nature Neuroscience*, 22(12), 2023–2028. <https://doi.org/10.1038/s41593-019-0516-y>
- Franceschini, A., Costantini, I., Pavone, F. S., & Silvestri, L. (2020). Dissecting neuronal activation on a brain-wide scale with immediate early genes. *Frontiers in Neuroscience*, 14, 569517. <https://doi.org/10.3389/fnins.2020.569517>
- Goodwin, N. L., Nilsson, S. R. O., Choong, J. J., & Golden, S. A. (2022). Toward the explainability, transparency, and universality of machine learning for behavioral classification in neuroscience. *Current Opinion in Neurobiology*, 73, 102544. <https://doi.org/10.1016/j.conb.2022.102544>
- Gururajan, A., Kos, A., & Lopez, J. P. (2018). Preclinical stress research: Where are we headed? An early career investigator's perspective. *Stress (Amsterdam, Netherlands)*, 21(5), 384–388. <https://doi.org/10.1080/10253890.2018.1446519>
- Haidl, T. K., Hedderich, D. M., Rosen, M., Kaiser, N., Seves, M., Lichtenstein, T., Penzel, N., Wenzel, J., Kambeitz-Illankovic, L., Ruef, A., Popovic, D., Schultze-Lutter, F., Chisholm, K., Upthegrove, R., Salokangas, R. K. R., Pantelis, C., Meisenzahl, E., Wood, S. J., Brambilla, P., ... Koutsouleris, N. (2023). The non-specific nature of mental health and structural brain outcomes following childhood trauma. *Psychological Medicine*, 53(3), 1005–1014. <https://doi.org/10.1017/S0033291721002439>
- Hamilton, P. J., Burek, D. J., Lombroso, S. I., Neve, R. L., Robison, A. J., Nestler, E. J., & Heller, E. A. (2018). Cell-type-specific epigenetic editing at the fosb gene controls susceptibility to social defeat stress. *Neuropsychopharmacology: official Publication of the American College of Neuropsychopharmacology*, 43(2), 272–284. <https://doi.org/10.1038/npp.2017.88>
- Häusel, A. S., Brix, L. M., Hartmann, J., Pöhlmann, M. L., Lopez, J.-P., Menegaz, D., Brivio, E., Engelhardt, C., Roeh, S., Bajaj, T., Rudolph, L., Stoffel, R., Hafner, K., Goss, H. M., Reul, J. M. H. M., Deussing, J. M., Eder, M., Ressler, K. J., Gassen, N. C., Chen, A., & Schmidt, M. V. (2021). The co-chaperone fkbp5 shapes the acute stress response in the paraventricular nucleus of the hypothalamus of male mice. *Molecular Psychiatry*, 26(7), 3060–3076. <https://doi.org/10.1038/s41380-021-01044-x>
- Hökfelt, T., Barde, S., Xu, Z. D., Kuteeva, E., Rüeegg, J., Le Maitre, E., Risling, M., Kehr, J., Ihnatko, R., Theodorsson, E., et al. (2018). Neuropeptide and small transmitter coexistence: Fundamental studies and relevance to mental illness. *Frontiers in Neural Circuits*, 12, 106. <https://doi.org/10.3389/fncir.2018.00106>
- Hsu, A. I., & Yttri, E. A. (2021). B-soid, an open-source unsupervised algorithm for identification and fast prediction of behaviors. *Nature Communications*, 12(1), 5188. <https://doi.org/10.1038/s41467-021-25420-x>
- Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Sanislow, C., & Wang, P. (2010). Research domain criteria (rdc): Toward a new classification framework for research on mental disorders. *The American Journal of Psychiatry*, 167(7), 748–751. <https://doi.org/10.1176/appi.ajp.2010.09091379>
- Isslser, O., van der Zee, Y. Y., Ramakrishnan, A., Wang, J., Tan, C., Loh, Y.-H. E., Purushothaman, I., Walker, D. M., Lorsch, Z. S., Hamilton, P. J., Peña, C. J., Flaherty, E., Hartley, B. J., Torres-Berrio, A., Parise, E. M., Kronman, H., Duffy, J. E., Estill, M. S., Calipari, E. S., ... Nestler, E. J. (2020). Sex-specific role for the long non-coding rna linc00473 in depression. *Neuron*, 106(6), 912–926.e915. <https://doi.org/10.1016/j.neuron.2020.03.023>
- Karamihalev, S., Brivio, E., Flachskamm, C., Stoffel, R., Schmidt, M. V., & Chen, A. (2020). Social dominance mediates behavioral adaptation to chronic stress in a sex-specific manner. *Elife*, 9, e58723. <https://doi.org/10.7554/eLife.58723>
- Kawashima, T., Kitamura, K., Suzuki, K., Nonaka, M., Kamijo, S., Takemoto-Kimura, S., Kano, M., Okuno, H., Ohki, K., & Bito, H. (2013). Functional labeling of neurons and their projections using the synthetic activity-dependent promoter e-sare. *Nature Methods*, 10(9), 889–895. <https://doi.org/10.1038/nmeth.2559>
- Kawashima, T., Okuno, H., & Bito, H. (2014). A new era for functional labeling of neurons: Activity-dependent promoters have come of age. *Frontiers in Neural Circuits*, 8, 37.
- Koutlas, I., Linders, L. E., van der Starre, S. E., Wolterink-Donselaar, I. G., Adan, R. A. H., & Meye, F. J. (2022). Characterizing and trapping a social stress-activated neuronal ensemble in the ventral tegmental area. *Frontiers in Behavioral Neuroscience*, 16, 936087. <https://doi.org/10.3389/fnbeh.2022.936087>
- Koutsouleris, V. W. (2022). Neurominer [computer software]. GitHub.
- Kuti, D., Winkler, Z., Horváth, K., Juhász, B., Szilvásy-Szabó, A., Fekete, C., Ferenczi, S., & Kovács, K. J. (2022). The metabolic stress response: Adaptation to acute-, repeated- and chronic challenges in mice. *iScience*, 25(8), 104693. <https://doi.org/10.1016/j.isci.2022.104693>
- Kwon, D. Y., Xu, B., Hu, P., Zhao, Y.-T., Beagan, J. A., Nofziger, J. H., Cui, Y., Phillips-Cremins, J. E., Blendy, J. A., Wu, H., & Zhou, Z. (2022). Neuronal yin yang1 in the prefrontal cortex regulates transcriptional

- and behavioral responses to chronic stress in mice. *Nature Communications*, 13(1), 55. <https://doi.org/10.1038/s41467-021-27571-3>
- Lauer, J., Zhou, M., Ye, S., Menegas, W., Schneider, S., Nath, T., Rahman, M. M., Di Santo, V., Soberanes, D., Feng, G., Murthy, V. N., Lauder, G., Dulac, C., Mathis, M. W., & Mathis, A. (2022). Multi-animal pose estimation, identification and tracking with deeplabcut. *Nature Methods*, 19(4), 496–504. <https://doi.org/10.1038/s41592-022-01443-0>
- Lin, R., Lopez, J. P., Cruceanu, C., Pierotti, C., Fiori, L. M., Squassina, A., Chillotti, C., Dieterich, C., Mellios, N., & Turecki, G. (2021). Circular rna circrcnt2 is upregulated in the anterior cingulate cortex of individuals with bipolar disorder. *Translational Psychiatry*, 11(1), 629. <https://doi.org/10.1038/s41398-021-01746-4>
- Lopez, J. P., Brivio, E., Santambrogio, A., De Donno, C., Kos, A., Peters, M., Rost, N., Czamara, D., Brückl, T. M., Roeh, S., et al. (2021). Single-cell molecular profiling of all three components of the hpa axis reveals adrenal abcb1 as a regulator of stress adaptation. *Science Advances*, 7(5), eabe4497. <https://doi.org/10.1126/sciadv.abe4497>
- Lopez, J. P., Kos, A., & Turecki, G. (2018). Major depression and its treatment: Micromars as peripheral biomarkers of diagnosis and treatment response. *Current Opinion in Psychiatry*, 31(1), 7–16. <https://doi.org/10.1097/YCO.0000000000000379>
- Lopez, J. P., Lücken, M. D., Brivio, E., Karamihalev, S., Kos, A., De Donno, C., Benjamin, A., Yang, H., Dick, A. L. W., Stoffel, R., Flachskamm, C., Ressler, A., Roeh, S., Huettl, R.-E., Parl, A., Eggert, C., Novak, B., Yan, Y., Yeoh, K., ... Chen, A. (2022). Ketamine exerts its sustained antidepressant effects via cell-type-specific regulation of kcnq2. *Neuron*, 110(14), 2283–2298.e2289. <https://doi.org/10.1016/j.neuron.2022.05.001>
- Luecken, M. D., & Theis, F. J. (2019). Current best practices in single-cell rna-seq analysis: A tutorial. *Molecular Systems Biology*, 15(6), e8746. <https://doi.org/10.15252/msb.20188746>
- Luotonen, S., Tikka, M., Karlsson, H., & Salokangas, R. K. (2013). Childhood trauma and distress experiences associate with psychotic symptoms in patients attending primary and psychiatric outpatient care. Results of the radeep study. *European Psychiatry: The Journal of the Association of European Psychiatrists*, 28(3), 154–160. <https://doi.org/10.1016/j.eurpsy.2011.11.005>
- Luxem, K., Mocellin, P., Fuhrmann, F., Kürsch, J., Miller, S. R., Palop, J. J., Remy, S., & Bauer, P. (2022). Identifying behavioral structure from deep variational embeddings of animal motion. *Communications Biology*, 5(1), 1267. <https://doi.org/10.1038/s42003-022-04080-7>
- Mandino, F., Cerri, D. H., Garin, C. M., Straathof, M., van Tilborg, G. A. F., Chakravarty, M. M., Dhenain, M., Dijkhuizen, R. M., Gozzi, A., Hess, A., et al. (2019). Animal functional magnetic resonance imaging: Trends and path toward standardization. *Frontiers in Neuroinformatics*, 13, 78.
- Marks, M., Qiuhan, J., Sturman, O., von Ziegler, L., Kollmorgen, S., von der Behrens, W., Mante, V., Bohacek, J., & Yanik, M. F. (2022). Deep-learning based identification, tracking, pose estimation, and behavior classification of interacting primates and mice in complex environments. *Nature Machine Intelligence*, 4(4), 331–340. <https://doi.org/10.1038/s42256-022-00477-5>
- Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). Deeplabcut: Markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, 21(9), 1281–1289. <https://doi.org/10.1038/s41593-018-0209-y>
- Maynard, K. R., Collado-Torres, L., Weber, L. M., Uyttingco, C., Barry, B. K., Williams, S. R., Catalini, J. L., Tran, M. N., Besich, Z., Tippani, M., Chew, J., Yin, Y., Kleinman, J. E., Hyde, T. M., Rao, N., Hicks, S. C., Martinowich, K., & Jaffe, A. E. (2021). Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. *Nature Neuroscience*, 24(3), 425–436. <https://doi.org/10.1038/s41593-020-00787-0>
- Miranda, L., Paul, R., Pütz, B., Koutsouleris, N., & Müller-Myhsok, B. (2021). Systematic review of functional mri applications for psychiatric disease subtyping. *Frontiers in Psychiatry*, 12, 665536. <https://doi.org/10.3389/fpsy.2021.665536>
- Moffitt, J. R., Lundberg, E., & Heyn, H. (2022). The emerging landscape of spatial profiling technologies. *Nature Reviews. Genetics*, 23(12), 741–759. <https://doi.org/10.1038/s41576-022-00515-3>
- Morris, S. E., Sanislow, C. A., Pacheco, J., Vaidyanathan, U., Gordon, J. A., & Cuthbert, B. N. (2022). Revisiting the seven pillars of rdcc. *BMC Medicine*, 20(1), 220. <https://doi.org/10.1186/s12916-022-02414-0>
- Musazzi, L., & Marrocco, J. (2016). The many faces of stress: Implications for neuropsychiatric disorders. *Neural Plasticity*, 2016, 8389737. <https://doi.org/10.1155/2016/8389737>
- Musazzi, L., Tornese, P., Sala, N., & Popoli, M. (2018). What acute stress protocols can tell us about ptsd and stress-related neuropsychiatric disorders. *Frontiers in Pharmacology*, 9, 758. <https://doi.org/10.3389/fphar.2018.00758>
- Nilsson, S. (2022). Simple behavioral analysis (simba) – an open source toolkit for computer classification of complex social behaviors in experimental animals. *BioRxiv*.
- Niu, M., Kasai, A., Tanuma, M., Seiriki, K., Igarashi, H., Kuwaki, T., Nagayasu, K., Miyaji, K., Ueno, H., Tanabe, W., et al. (2022). Claustrum mediates bidirectional and reversible control of stress-induced anxiety responses. *Science Advances*, 8(11), eabi6375. <https://doi.org/10.1126/sciadv.abi6375>
- Palla, G., Spitzer, H., Klein, M., Fischer, D., Schaar, A. C., Kuemmerle, L. B., Rybakov, S., Ibarra, I. L., Holmberg, O., Virshup, I., Lotfollahi, M., Richter, S., & Theis, F. J. (2022). Squidpy: A scalable framework for spatial omics analysis. *Nature Methods*, 19(2), 171–178. <https://doi.org/10.1038/s41592-021-01358-2>
- Pereira, T. D., Tabris, N., Matsliah, A., Turner, D. M., Li, J., Ravindranath, S., Papadoyannis, E. S., Normand, E., Deutsch, D. S., Wang, Z. Y., McKenzie-Smith, G. C., Mitelut, C. C., Castro, M. D., D’Uva, J., Kislin, M., Sanes, D. H., Kocher, S. D., Wang, S. S.-H., Falkner, A. L., Shaevitz, J. W., & Murthy, M. (2022). Slep: A deep learning system for multi-animal pose tracking. *Nature Methods*, 19(4), 486–495. <https://doi.org/10.1038/s41592-022-01426-1>
- Popovic, D., Ruef, A., Dwyer, D. B., Antonucci, L. A., Eder, J., Sanfelici, R., Kambeitz-Illankovic, L., Oztuerk, O. F., Dong, M. S., Paul, R., Paolini, M., Hedderich, D., Haidl, T., Kambeitz, J., Ruhmann, S., Chisholm, K., Schultze-Lutter, F., Falkai, P., Pergola, G., ... Koutsouleris, N., PRONIA Consortium. (2020). Traces of trauma: A multivariate pattern analysis of childhood trauma, brain structure, and clinical phenotypes. *Biological Psychiatry*, 88(11), 829–842. <https://doi.org/10.1016/j.biopsych.2020.05.020>
- Ramirez, S., Liu, X., MacDonald, C. J., Moffa, A., Zhou, J., Redondo, R. L., & Tonegawa, S. (2015). Activating positive memory engrams suppresses depression-like behaviour. *Nature*, 522(7556), 335–339. <https://doi.org/10.1038/nature14514>
- Renier, N., Wu, Z., Simon, D. J., Yang, J., Ariel, P., & Tessier-Lavigne, M. (2014). Idisco: A simple, rapid method to immunolabel large tissue samples for volume imaging. *Cell*, 159(4), 896–910. <https://doi.org/10.1016/j.cell.2014.10.010>
- Russell, G., & Lightman, S. (2019). The human stress response. *Nature Reviews. Endocrinology*, 15(9), 525–534. <https://doi.org/10.1038/s41574-019-0228-0>
- Sanacora, G., Yan, Z., & Popoli, M. (2022). The stressed synapse 2.0: Pathophysiological mechanisms in stress-related neuropsychiatric disorders. *Nature Reviews. Neuroscience*, 23(2), 86–103. <https://doi.org/10.1038/s41583-021-00540-x>
- Satija, R., Farrell, J. A., Gennert, D., Schier, A. F., & Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. *Nature Biotechnology*, 33(5), 495–502. <https://doi.org/10.1038/nbt.3192>
- Scharf, S. H., Liebl, C., Binder, E. B., Schmidt, M. V., & Müller, M. B. (2011). Expression and regulation of the fkbp5 gene in the adult mouse brain. *PLoS One*, 6(2), e16883. <https://doi.org/10.1371/journal.pone.0016883>
- Schneider, S., Lee, J. H., & Mathis, M. W. (2022). Learnable latent embeddings for joint behavioral and neural analysis. *ArXiv*.
- Segalin, C., Williams, J., Karigo, T., Hui, M., Zelikowsky, M., Sun, J. J., Perona, P., Anderson, D. J., & Kennedy, A. (2021). The mouse action recognition system (mars) software pipeline for automated analysis of social behaviors in mice. *Elife*, 10, e63720. <https://doi.org/10.7554/eLife.63720>
- Shemesh, Y., & Chen, A. (2023). A paradigm shift in translational psychiatry through rodent neuroethology. *Molecular Psychiatry*, 28, 993–1003. <https://doi.org/10.1038/s41380-022-01913-z>

- Shemesh, Y., Forkosh, O., Mahn, M., Anpilov, S., Sztainberg, Y., Manashirov, S., Shlapobersky, T., Elliott, E., Tabouy, L., Ezra, G., Adler, E. S., Ben-Efraim, Y. J., Gil, S., Kuperman, Y., Haramati, S., Dine, J., Eder, M., Deussing, J. M., Schneidman, E., Yizhar, O., & Chen, A. (2016). Ucn3 and crf-r2 in the medial amygdala regulate complex social dynamics. *Nature Neuroscience*, 19(11), 1489–1496. <https://doi.org/10.1038/nn.4346>
- Shemesh, Y., Sztainberg, Y., Forkosh, O., Shlapobersky, T., Chen, A., & Schneidman, E. (2013). High-order social interactions in groups of mice. *eLife*, 2, e00759. <https://doi.org/10.7554/eLife.00759>
- Short, A. K. (2021). *Single-cell transcriptional changes in hypothalamic corticotropin-releasing factor-expressing neurons after early-life adversity inform enduring alterations in vulnerabilities to stress* (T. Z. Baram, Ed.) Biological Psychiatry Global Open Science.
- Silva, B. A., Burns, A. M., & Gräff, J. (2019). A cfos activation map of remote fear memory attenuation. *Psychopharmacology*, 236(1), 369–381. <https://doi.org/10.1007/s00213-018-5000-y>
- Stahlschmidt, S. R., Ulfenborg, B., & Synnergren, J. (2022). Multimodal deep learning for biomedical data fusion: A review. *Brief Bioinform*, 23(2), 1–15.
- Sturman, O., von Ziegler, L., Schläppi, C., Akyol, F., Privitera, M., Slominski, D., Grimm, C., Thieren, L., Zerbi, V., Grewe, B., & Bohacek, J. (2020). Deep learning-based behavioral analysis reaches human accuracy and is capable of outperforming commercial solutions. *Neuropsychopharmacology: official Publication of the American College of Neuropsychopharmacology*, 45(11), 1942–1952. <https://doi.org/10.1038/s41386-020-0776-y>
- Tian, L., Chen, F., & Macosko, E. Z. (2022). The expanding vistas of spatial transcriptomics. *Nature Biotechnology*. <https://doi.org/10.1038/s41587-022-01448-2>
- Tighe, P. J., Ryder, R. R., Todd, I., & Fairclough, L. C. (2015). Elisa in the multiplex era: Potentials and pitfalls. *Proteomics. Clinical Applications*, 9(3-4), 406–422. <https://doi.org/10.1002/prca.201400130>
- Velten, B., Braunger, J. M., Argelaguet, R., Arnol, D., Wirbel, J., Bredikhin, D., Zeller, G., & Stegle, O. (2022). Identifying temporal and spatial patterns of variation from multimodal data using mefisto. *Nature Methods*, 19(2), 179–186. <https://doi.org/10.1038/s41592-021-01343-9>
- von Mücke-Heim, I. A., Urbina-Treviño, L., Bordes, J., Ries, C., Schmidt, M. V., & Deussing, J. M. (2022). Introducing a depression-like syndrome for translational neuropsychiatry: A plea for taxonomical validity and improved comparability between humans and mice. *Molecular Psychiatry*. 28(1), 329–340. <https://doi.org/10.1038/s41380-022-01762-w>
- von Ziegler, L. M., Floriou-Servou, A., Waag, R., Das Gupta, R. R., Sturman, O., Gapp, K., Maat, C. A., Kockmann, T., Lin, H.-Y., Duss, S. N., Privitera, M., Hinte, L., von Meyenn, F., Zeilhofer, H. U., Germain, P.-L., & Bohacek, J. (2022). Multiomic profiling of the acute stress response in the mouse hippocampus. *Nature Communications*, 13(1), 1824. <https://doi.org/10.1038/s41467-022-29367-5>
- Wang, X.-J., Hu, H., Huang, C., Kennedy, H., Li, C. T., Logothetis, N., Lu, Z.-L., Luo, Q., Poo, M.-M., Tsao, D., Wu, S., Wu, Z., Zhang, X., & Zhou, D. (2020). Computational neuroscience: A frontier of the 21. *National Science Review*, 7(9), 1418–1422. <https://doi.org/10.1093/nsr/nwaa129>
- Wen, L., Li, G., Huang, T., Geng, W., Pei, H., Yang, J., Zhu, M., Zhang, P., Hou, R., Tian, G., Su, W., Chen, J., Zhang, D., Zhu, P., Zhang, W., Zhang, X., Zhang, N., Zhao, Y., Cao, X., ... Chen, Z.-J. (2022). Single-cell technologies: From research to application. *Innovation (Cambridge (Mass.))*, 3(6), 100342. <https://doi.org/10.1016/j.xinn.2022.100342>
- Wolf, F. A., Angerer, P., & Theis, F. J. (2018). Scanpy: Large-scale single-cell gene expression data analysis. *Genome Biology*, 19(1), 15. <https://doi.org/10.1186/s13059-017-1382-0>
- Wu, Z., Lin, D., & Li, Y. (2022). Pushing the frontiers: Tools for monitoring neurotransmitters and neuromodulators. *Nature Reviews. Neuroscience*, 23(5), 257–274. <https://doi.org/10.1038/s41583-022-00577-6>
- Ye, S., Mathis, A., & Mathis, M. W. (2022). Panoptic animal pose estimators are zero-shot performers. *Arxiv*,

Chapter 8

General discussion

8.1 Summary

The primary objective of this thesis was to examine the precise role of FKBP51 within specific stress-related brain regions on the underlying neurobiological mechanisms and behavioral profiles associated with stress-related disorders. However, the conventional preclinical tools used to assess behavioral symptoms, such as social aversion, are not sufficient and do not capture social behavioral domains in detail due to their reliance on reductionistic approaches. These approaches involve testing animals in controlled environments that often fail to reflect ethologically relevant behavioral constructs. Recent advancements in automatically annotated behavioral assessment using high-throughput pose estimation tools [162, 163] have opened up possibilities for more comprehensive behavioral analyses. This thesis investigated various automatically annotated behavioral assessment tools for preclinical psychiatry research.

In chapter 2, the different social behavioral assessment methods throughout history were reviewed, aiming to explore novel strategies for deep behavioral phenotyping using automated annotated behavioral assessment tools, particularly in the field of psychiatry research. The review demonstrates how these tools can combine strengths from the fields of ethology and comparative psychology, leading to the development of a novel social behavioral task. This task incorporates a semi-naturalistic setup that enables the expression of more natural behaviors while maintaining control over the environment by restricting space and external influences.

In chapter 3 the open-source python package, called “DeepOF” is introduced to investigate the behavioral profile of rodents using advanced computational tools. DeepOF utilizes high-throughput pose estimation tools, leveraging DeepLabCut annotated data, and enables post-hoc analysis through both supervised classification and unsupervised clustering. The supervised classification is intended to extract pre-defined and characterized traits, whereas

unsupervised clustering aims to explore the data and extract patterns without external information. The package allows for exploration of individual and social behavioral profiles in rodents using the supervised classification, as explored in chapter 4 and chapter 6. In addition, DeepOF can be further customized into the behavioral phenotyping of any required behavioral analysis using an unsupervised analysis, as explored on the social behavioral profile in chapter 4, or on fear-related behavior in chapter 5. DeepOF contributes to the standardization and reproducibility of the different behavioral constructs by providing an automated and validated open-source tool that is uniquely optimized for top-down video recordings.

In chapter 4 the established deep phenotyping strategies of DeepOF are employed using supervised classification and unsupervised clustering to investigate the individual and social behavioral profiles following chronic social defeat stress. The application of DeepOF's pipelines revealed a distinct stress-induced social behavioral pattern, particularly evident during initial encounters and diminishing over time due to habituation. In addition, while the classical social avoidance task did identify the stress-induced social behavioral differences, both DeepOF behavioral pipelines provided a clearer and more detailed profile. Ultimately, DeepOF enhances the classification of affected individual and social behaviors in stress-related disorders, potentially aiding in drug development for psychiatric disorders.

In chapter 5 the sex-specific effects of early life stress exposure were investigated on HPA-axis dysregulation. The chapter examines alterations of *Fkbp5* expression in several stress-related brain regions including the BLA and subregions of the dorsal hippocampus, but also examines the adrenal weight, and baseline CORT levels in a sex- and time-dependent manner. Fear-related behaviors were assessed using classical freezing behavioral analysis as well as in-depth behavioral analysis using DeepOF's unsupervised clustering analysis. The passive fear behavioral response, as identified via freezing behavior, and also the active fear response, as identified using an unsupervised analysis, were altered by ELS exposure in a sex-specific manner. The DeepOF unsupervised analysis is highly customizable for different behavioral setups and analysis protocols and provides an additional layer to explore fear-related behaviors without prior assumptions. DeepOF unsupervised analysis allows for a hypothesis-generating behavioral analysis, which ultimately can lead to a better understanding of the stress-induced behavioral phenotype. The findings demonstrated that early life stress exposure alters both passive and active fear responses in a sex-specific manner.

In chapter 6 the role of *Fkbp5* expression in the LC was investigated concerning different stress exposures and its impact on the social behavioral profile using DeepOF's supervised classification strategies. Only acute social stress led to a significant upregulation of *Fkbp5* mRNA, highlighting the importance of the social nature of stressors in *Fkbp5* regulation within the LC. To comprehend *Fkbp5* regulation within the LC, a conditional *Fkbp5* knockout line within the noradrenergic system (*Fkbp5^{Nat}*) was generated. Notably, male *Fkbp5^{Nat}* mice exhibited altered social behavioral profile towards a novel social conspecific, with no significant effects observed in female mice. Furthermore, *Fkbp5^{Nat}* mice exhibited long-lasting changes in the molecular associated with to the noradrenergic synapse under baseline conditions and demonstrated modified noradrenergic signaling when exposed to a novel social stimulus, specifically in the BLA.

In chapter 7, the discussion section begins by emphasizing the significance of two essential pillars that will shape the future of stress-related research. These pillars involve the advancements made in molecular neurobiology and computational neuroscience, which have contributed to a better understanding of the intricate neurobiological pathways and complex behaviors underlying stress. Importantly, both fields have seen a rapid development of novel techniques, but they are also strongly interdependent on each other. For instance, the success of novel molecular neurobiological techniques like single-cell sequencing relies heavily on the analytical tools developed by computational neuroscience.

To summarize, this thesis introduces innovative deep phenotyping strategies that utilize automated motion tracking data to investigate rodent behavior in relation to the effects of genetic factors (*FKBP5*) and environmental influences (stress exposure). The thesis underlines the brain-region specific effects of stress exposure and *Fkbp5* regulation on the social and fear-related behavioral phenotypes.

8.2 Next generation deep phenotyping requires more than motion tracking data

In recent years, there has been a growing focus on deep phenotyping of behavior using motion tracking data, facilitated by advancements in open-source computational tools for pose estimation [162–167], and their downstream analysis tools, including supervised classification

[167–170], and unsupervised clustering [170–175]. Understanding the behavioral adaptation to genetic- and environmental factors is crucial. To achieve a more comprehensive understanding of these factors on behavior, it is essential to investigate the underlying neurobiological mechanisms that contribute to these behavioral changes and to combine the advances in behavioral analysis with deep phenotyping of additional physiological and functional domains.

On the level of inter-cellular signalling, exposure to a stressful event triggers the release of a specific set of stress mediator molecules, including monoamines, neuropeptides, and steroid hormones, which transmit the stress signal to specific brain regions and contribute to functional changes in the brain [34]. In-vivo measurements of these stress mediators provide additional insights into the effects of stress exposure alongside the behavioral phenotype. An established method to investigate the release of stress mediators in-vivo within the brain is microdialysis. This technique relies on the principle of diffusion, where molecules pass through a semipermeable membrane into an artificial cerebrospinal fluid and enables the measurement of quantifiable concentrations of the molecule of interest within determined time intervals. Microdialysis allows for recordings that can span hourly or even overnight durations, which is particularly important when investigating fluctuations of stress mediators influenced by the circadian rhythm, as recently highlighted again by Upton et al. 2023 [176]. This method enables the investigation of the underlying neurobiological pathway and can provide valuable insights into the impact of genetic factors, such as *Fkbp5*, on psychiatric risk exposure. This was illustrated by Anderzhanova et al., 2020 [177] that showed the significant role of *Fkbp5* regulation in the changes induced by ketamine in the release of mature Brain-Derived Neurotrophic Factor within the medial prefrontal cortex. Furthermore, in chapter 6, we present findings that highlight the influence of *Fkbp5* regulation in the noradrenergic system on the release of NE within the BLA, particularly following social stimulation. However, the time resolution of microdialysis is limited, as stable measurements require a minimum sampling time of around 10-20 minutes. As a result, in chapter 6, it was only possible to compare baseline measurements with measurements taken after social stimulation. Enhancing the time resolution would enable a direct comparison of stress mediator release, such as NE, with advanced behavioral analysis profiles, as performed by DeepOF or other behavioral analysis packages. The emergence of genetically encoded GPCR activation-based (GRAB) sensors represents a significant advancement, as they exhibit an exceptional level of time resolution,

capable of detecting changes in stress mediators within the millisecond time frame [178]. This high temporal resolution makes GRAB sensors a powerful tool for investigating the cellular response to stress in the brain, as discussed in more detail in chapter 7. In the context of chapter 6, the utilization of GRAB sensors would enable a direct comparison of fluctuations in NE signaling within the BLA with the supervised behavioral classifiers identified by DeepOF, as well as potentially with unsupervised behavioral clusters. By incorporating GRAB sensors, the integration of additional data modalities becomes possible, thereby expanding the scope of deep phenotyping in mice using motion tracking data. It is worth noting that various other in-vivo data modalities can also be integrated with motion tracking data, as discussed in chapter 2.

Another intriguing advancement in adding data modalities to behavior is the in-vivo recording of neuronal activity, which in recent years has enabled the simultaneous monitoring of hundreds to thousands of neurons using state-of-the-art technologies, such as neuropixels [179–182]. The analysis pipeline for such experiments has become increasingly complex, especially when combined with behavioral data. Incorporating multiple data modalities inevitably leads to a larger volume of data to be analyzed, presenting a challenge in terms of data integration. However, a recent groundbreaking study introduced an open-source analysis pipeline called “CEBRA,” which leverages the patterns of neuronal activity from hundreds of neurons to decode underlying behavior [183]. CEBRA enables the integration of these data modalities, allowing for a more comprehensive supervised and unsupervised representation of behavior through combined embeddings of motion tracking and neural activity data [183].

In conclusion, the methodological advancements in neuronal tracking tools such as GRAB sensors and neuropixels offer valuable opportunities to incorporate additional data modalities in behavioral phenotyping. While the increased information poses challenges in data integration, recent open-source analysis tools, such as CEBRA demonstrate the potential and facilitate a deeper understanding of the underlying mechanisms associated with the behavior of interest.

8.3 Decoding the ambiguous nature of FKBP51 on psychiatric risk through brain region and cell type specific effects

The regulation of *Fkbp5* and its impact on stress vulnerability is a complex and nuanced phenomenon that is mediated in a brain and cell type specific manner. Initially, it was discovered that certain polymorphisms in the FKBP5 gene led to higher expression patterns, increasing susceptibility to stress-related disorders [97]. However, the heightened response to antidepressant treatment in risk-allele carriers suggests that the regulation of FKBP5 and its impact on stress vulnerability involve intricate mechanisms that warranted further investigation.

The expression pattern of *Fkbp5* in response to different stress paradigms is highly specific. Exposure to different acute stressors elicits brain region specific increases in *Fkbp5* expression, where for instance restraint stress affects the PVN and central amygdala, while food deprivation affects the hippocampus [109]. The exposure to chronic stress increased the expression of *Fkbp5* throughout various stress-related brain regions, including the nucleus accumbens, hippocampus, amygdala, and prefrontal cortex [111, 112], but not the PVN [184]. The present thesis delves deeper into the regulation of *Fkbp5* under different stress paradigms. It is revealed that chronic early life stress exposure specifically increases *Fkbp5* expression in the CA1 region of the dorsal hippocampus in male mice (see chapter 5). Furthermore, in chapter 6 it is demonstrated that *Fkbp5* expression in the LC is only increased after acute social defeat stress, not restraint stress or a history of chronic social defeat stress. The dorsal raphe nucleus follows a different pattern of *Fkbp5* regulation, in which *Fkbp5* is upregulated after acute social defeat stress and chronic social defeat stress. These findings contribute to the further disentanglement of *Fkbp5*'s role in different stress paradigms and highlight the significance of brain region specific regulation of the *Fkbp5* gene.

The importance of *Fkbp5* regulation between different brain regions on downstream functional and behavioral outcomes was further illustrated by Engelhardt et al., 2021 [185]. They showed a disrupted HPA-axis function and anxiogenic behavioral phenotypes in animals lacking *Fkbp5* in the BNST, while mice with *Fkbp5* overexpression in the BNST exhibited a protected phenotype. These findings reveal a complex and ambiguous role of *Fkbp5* regulation, where divergent patterns of regulation within the BNST play a critical role in de-

terminating the response to stress exposure. This is in contrast with earlier findings on full *Fkbp5* knockout mice, which displayed a more resilient phenotype after stress exposure by active stress-coping behavior [115, 117], therefore conforming the region specific role of *Fkbp5* regulation. Further emphasizing the brain region and cell type specific effects of *Fkbp5*, van Doeselaar et al. (2023) [186] demonstrated opposing effects on behavior, brain structure, and gene expression profiles depending on the cell type specific knockout of *Fkbp5* in GABAergic or glutamatergic forebrain neurons in a sex-dependent manner. The current thesis expands on the knowledge of the brain region and cell type specific effects of *Fkbp5* regulation. In chapter 6, the conditional knock-out of *Fkbp5* within the noradrenergic system induced a prosocial behavioral phenotype only in male mice. Additionally, these mice exhibited altered molecular pathways related to the noradrenergic synapse under baseline conditions and show altered noradrenergic signaling when exposed to a novel social stimulus. In conclusion, the regulation of *Fkbp5* and its effect on stress vulnerability is strongly dependent on the context of stress exposure. The collective data from this thesis together with previously published observations underline that the role of *Fkbp5* on stress exposure is ambiguous and depends on brain region and cell type specific regulation of *Fkbp5*. These findings have significant implications for enhancing our understanding of the role of the psychiatric risk factor FKBP51 in relation to the vulnerability of stress-related disorders. Moreover, these findings contribute to unraveling the complex neurobiological mechanisms underlying FKBP51 signaling. Ultimately, by better understanding the intricate neurobiological pathways related to FKBP51 signaling, future research may be able to identify more precise targets for intervention, leading to more effective treatments for individuals affected by these disorders.

The outcome of stress exposure is not solely determined by *Fkbp5*-related gene expression changes in different brain regions and cell types, but are also influenced by various other factors. Notably, the effects of *Fkbp5* knockout in GABAergic and glutamatergic forebrain neurons [186], as well as noradrenergic cells (see chapter 6), are strongly modulated by sex. Hence, the inclusion of factors such as sex is imperative for a comprehensive understanding of the intricate role of *Fkbp5* in the stress response system.

8.4 The influence of sex on stress vulnerability: a plea for more ethologically relevant stressors

The outcome of stress exposure is determined by a combination of genetic-, epigenetic-, and environmental factors. This thesis highlights the critical role of the combined effects of these factors on stress exposure and discusses the particular important role of advanced tracking tools in integrating these diverse factor on the outcome of the behavioral profile, see chapter 2. For instance, it has long been recognized that female rodents exhibit a higher physiological stress response compared to males, as evidenced by elevated CORT levels following exposure to various stressors [187]. More recent findings have further revealed sex differences in CRF signaling within LC neurons, with females showing greater sensitivity and higher CRF levels compared to males, making the LC an intriguing target for investigating stress-induced sex differences. [188, 189]. The influence of sex on stress vulnerability has gained considerable attention in recent years, with an increased amount of evidence supporting sex-dependent regulation of stress-related behavioral symptoms across different stress paradigms [186, 190, 191].

To compare stress-induced behavioral effects between sexes, it is crucial to employ comparable stress paradigms that account for inherent sex differences. The innate differences between sexes need to be recognized in order to maintain the observation of ethologically relevant behaviors. This is particularly relevant when examining social behavioral patterns, which are heavily influenced by sex-related differences, for instance as observed in hierarchy, parental care, and sexual behaviors [192]. An interesting example pertains to adapting the traditionally male-focused chronic social defeat stress paradigm for females [193–195]. This paradigm typically involves the exposure of male mice to an aggressive male mouse, usually from a different strain, that subsequently attacks and forces the experimental mouse into subordination [196]. While this stress paradigm is effective in male mice due to their expression of strong aggressive behaviors in the context of hierarchy towards other males, it is not optimal for females, as the same extent of aggression is not observed among females [197]. Consequently, the majority of social defeat studies utilize a male aggressor mouse targeting another male. Recent contributions have adapted the social stress protocol by tricking the male aggressor mouse by employing male urine on the experimental female mouse, leading

to robust attacks towards the female mice [193, 194]. An important symptom of the chronic social defeat stress model in males is the increase in social avoidance behavior in stressed mice [196, 198]. However, female stressed mice did not exhibit a difference in this particular social avoidance behavior. Additionally, a recent study conducted by Pantoja et al., 2023 [195], utilized an accelerated version of the social defeat stress model for female mice. The study employed a novel approach where the aggressor mice were initially exposed to a male mouse for 30 seconds, followed by swapping the male mouse with a female experimental mouse, which was then subjected to attacks. The findings from this study align with the observations made in the study by van Doeselaar et al., 2021 [194], in which female mice did not exhibit an increase in social avoidance behavior. Additionally, Pantoja et al., 2023 [195] specifically compared the distribution of social avoidance susceptibility versus resiliency between males and females. The results revealed a higher percentage of resilient mice among females (85%) compared to males (55%). However, it is noteworthy that the number of attacks experienced by females was lower than that of males throughout all sessions of the defeat model. In the study conducted by van Doeselaar et al. (2021) [194], although the specific number of attacks was not measured, it was observed that female mice experienced attacks in only 60% of the cases. However, a direct comparison with male mice was not conducted in the study. These studies suggest that there are differences in the frequency and severity of attacks between males and females subjected to attacks from male mice aggressors, which could potentially contribute to variations in social avoidance behavior between the two sexes. Furthermore, the social defeat stress paradigm in females is far from the ethological relevant behaviors that are normally exhibited towards female mice, which could further explain the differences in social avoidance behavior between sexes. Female mice may exhibit different responses to stressors that better align with their ethological context.

Another option for chronic social stress exposure across sexes is the social instability stress paradigm, where group composition is changed multiple times over a prolonged period [199]. However, the inherent differences in aggressive behaviors within hierarchical contexts between males and females remain a challenge in this paradigm, as males will exhibit more aggressive behavior among each other than females [200]. Lopez et al., 2021 [201] emphasize sex-specific differences in stress sensitivity depending on the type of stressor, for instance as exhibited by a heightened sensitivity of females to social isolation. Therefore, instead of adapting stress

paradigms validated in male mice for use in females, a more promising approach may involve developing specific stress models that capture ethologically relevant aspects in females.

The integration of social stress exposure within a semi-naturalistic environment could offer a valuable method for inducing ethologically relevant stressors that are comparable between sexes. This approach would enable the investigation of group behaviors in a free interaction environment while maintaining control over the environment and space, as illustrated by [202, 203]. Different type of stressors can be added into such an environment, for instance via limiting food resources, changing the ambient temperature, or socially isolating the animals. Another avenue worth exploring is social stress transmission, where a single animal is exposed to a stressor within or outside the environment, and subsequent behavioral responses of other animals towards the previously stressed individual are examined.

The growing recognition of the importance of including both males and females in stress research is a promising and valuable advancement, especially considering the influence of sex on the prevalence of stress-related disorders [204]. While current efforts to adapt traditionally male-biased stress paradigms to include females are commendable, it is important to acknowledge that these adaptations may not fully capture the ethologically relevant behaviors for females in certain stress paradigms. Therefore, it is crucial to continue developing specific stress paradigms tailored to females in order to enhance our understanding of stress exposure across sexes. By doing so, we can gain more comprehensive insights into the impact of stress on both males and females and improve the translatability of findings to clinical applications.

8.5 Closing remarks

In conclusion, this thesis explored the brain region and cell type specific role of FKBP51 in stress-related disorders, by investigating the underlying neurobiological mechanisms and behavioral profile. The behavioral phenotyping is performed using advanced computational tools using automated motion tracking data to explore the behavioral profile using supervised behavioral classification and unsupervised clustering tools with the DeepOF package. The findings highlight the importance of deep phenotyping strategies for understanding genetic and environmental factors that influence behavior. Furthermore the need for integrating additional data modalities is emphasized, such as in-vivo measurements of stress mediators and neuronal activity, in order to gain a more comprehensive understanding of the behavioral

phenotype. This research contributes to our understanding of stress-related disorders and the intricate interplay between genetic factors, environmental influences, and the underlying neurobiological mechanisms. It provides valuable insights into the development of future research strategies for deep phenotyping and decoding the ambiguous nature of FKBP51 on psychiatric risk through brain- and cell-type specific effects.

Bibliography

- [1] W. H. O. Depression and other common mental disorders: global health estimates. Technical report, World Health Organization, 2017.
- [2] Theo Vos, Stephen S Lim, Cristiana Abbafati, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019. *The Lancet*, 396(10258):1204–1222, 2020.
- [3] Aleksandra Kupferberg, Lucy Bicks, and Gregor Hasler. Social functioning in major depressive disorder. *Neuroscience & Biobehavioral Reviews*, 69:313–332, 2016.
- [4] Stefano Porcelli, Nic Van Der Wee, Steven van der Werff, et al. Social brain, social dysfunction and social withdrawal. *Neuroscience & Biobehavioral Reviews*, 97:10–33, 2019.
- [5] Sebastian Trautmann, Jürgen Rehm, and Hans-Ulrich Wittchen. The economic costs of mental disorders: Do our societies react appropriately to the burden of mental disorders? *EMBO Reports*, 17(9):1245–1249, 2016.
- [6] Christine Heim and Charles B Nemeroff. The role of childhood trauma in the neurobiology of mood and anxiety disorders: preclinical and clinical studies. *Biological psychiatry*, 49(12):1023–1039, 2001.
- [7] Ronald C Kessler, Katie A McLaughlin, Jennifer Greif Green, et al. Childhood adversities and adult psychopathology in the who world mental health surveys. *The British Journal of Psychiatry*, 197(5):378–385, 2010.
- [8] L Mandelli, C Petrelli, and A Serretti. The role of specific early trauma in adult depression: A meta-analysis of published literature. childhood trauma and adult depression. *European Psychiatry*, 30(6):665–680, 2015.
- [9] Charles B Nemeroff. Paradise lost: the neurobiological and clinical consequences of child abuse and neglect. *Neuron*, 89(5):892–909, 2016.
- [10] Mirko Manchia, Anouk W Gathier, Hale Yapici-Eser, et al. The impact of the prolonged covid-19 pandemic on stress resilience and mental health: A critical review across waves. *European Neuropsychopharmacology*, 55:22–83, 2022.
- [11] Thole H Hoppen, Stefan Priebe, Inja Vetter, and Nexhmedin Morina. Global burden of post-traumatic stress disorder and major depression in countries affected by war between 1989 and 2019: a systematic review and meta-analysis. *BMJ global health*, 6(7):e006303, 2021.
- [12] Damian F Santomauro, Ana M Mantilla Herrera, Jamileh Shadid, et al. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the covid-19 pandemic. *The Lancet*, 398(10312):1700–1712, 2021.

- [13] George I Papakostas and Maurizio Fava. Does the probability of receiving placebo influence clinical trial outcome? a meta-regression of double-blind, randomized clinical trials in mdd. *European Neuropsychopharmacology*, 19(1):34–40, 2009.
- [14] Evelinda Trindade, Devidas Menon, Leigh-Ann Topfer, and Carlos Coloma. Adverse effects associated with selective serotonin reuptake inhibitors and tricyclic antidepressants: a meta-analysis. *CMAJ*, 159(10):1245–1252, 1998.
- [15] Pierre M Bet, Jacqueline G Hugtenburg, Brenda WJH Penninx, and Witte JG Hoogendijk. Side effects of antidepressants during long-term use in a naturalistic setting. *European Neuropsychopharmacology*, 23(11):1443–1451, 2013.
- [16] Koustuv Saha, John Torous, Emre Kiciman, et al. Understanding side effects of antidepressants: large-scale longitudinal study on social media data. *JMIR Mental Health*, 8(3):e26589, 2021.
- [17] Masaru Mimura. Comorbidity of depression and other diseases. *Japan Medical Association Journal*, 44(5):225–229, 2001.
- [18] Yemi Aina and Jeffrey L Susman. Understanding comorbidity with depression and anxiety disorders. *The Journal of the American Osteopathic Association*, 106(5 Suppl 2):S9–14, 2006.
- [19] David Goldberg. The heterogeneity of “major depression”. *World Psychiatry*, 10(3):226, 2011.
- [20] American Psychiatric Association. *Diagnostic and statistical manual of mental disorders: DSM-5*, volume 5. American Psychiatric Association Washington, DC, 2013.
- [21] Uma Rao. Dsm-5: disruptive mood dysregulation disorder. *Asian Journal of Psychiatry*, 11:119–123, 2014.
- [22] Yuri Milaneschi, Femke Lamers, Michael Berk, and Brenda WJH Penninx. Depression heterogeneity and its biological underpinnings: toward immunometabolic depression. *Biological Psychiatry*, 88(5):369–380, 2020.
- [23] Bruce S McEwen. Stress, adaptation, and disease: Allostasis and allostatic load. *Annals of the New York Academy of Sciences*, 840(1):33–44, 1998.
- [24] Bruce S McEwen. Physiology and neurobiology of stress and adaptation: central role of the brain. *Physiological Reviews*, 87(3):873–904, 2007.
- [25] Irina Milisav, Borut Poljsak, and Dušan Šuput. Adaptive response, evidence of cross-resistance and its potential clinical use. *International Journal of Molecular Sciences*, 13(9):10771–10806, 2012.
- [26] John A Updegraff and Shelley E Taylor. From vulnerability to growth: Positive and negative effects of stressful life events. In *Loss and trauma*, pages 3–28. Routledge, 2021.
- [27] C Bernard. La constitution physico-chimique du milieu intérieur [the physicochemical constitution of the environment within]. *Revue Scientifique*, 2:670–672, 1872.

- [28] Walter Bradford Cannon. Homeostasis. *The wisdom of the body*. Norton, Newyork, 1932.
- [29] Hans Selye. A syndrome produced by diverse nocuous agents. *Nature*, 138(3479):32–32, 1936.
- [30] Bruce S McEwen and Eliot Stellar. Stress and the individual: Mechanisms leading to disease. *Archives of Internal Medicine*, 153(18):2093–2101, 1993.
- [31] Peter Sterling. Allostasis: a new paradigm to explain arousal pathology. *Handbook on Life Stress, Cognition, and Health*, pages 629–649, 1990.
- [32] Bruce S McEwen. Allostasis and allostatic load: implications for neuropsychopharmacology. *Neuropsychopharmacology*, 22(2):108–124, 2000.
- [33] Gal Richter-Levin and Carmen Sandi. “labels matter: Is it stress or is it trauma?”. *Translational Psychiatry*, 11(1):385, 2021.
- [34] Marian Joëls and Tallie Z Baram. The neuro-symphony of stress. *Nature Reviews Neuroscience*, 10(6):459–466, 2009.
- [35] Lucas Miranda, Joeri Bordes, Serena Gasperoni, and Juan Pablo Lopez. Increasing resolution in stress neurobiology: from single cells to complex group behaviors. *Stress*, 26(1):2186141, 2023.
- [36] Lívea Dornela Godoy, Matheus Teixeira Rossignoli, Polianna Delfino-Pereira, et al. A comprehensive overview on stress neurobiology: basic concepts and clinical implications. *Frontiers in Behavioral Neuroscience*, 12:127, 2018.
- [37] Laurie Kelly McCorry. Physiology of the autonomic nervous system. *American Journal of Pharmaceutical Education*, 71(4), 2007.
- [38] Tyler LeBouef, Zachary Yaker, and Lacey Whited. Physiology, autonomic nervous system. In *StatPearls*. StatPearls Publishing, 2021.
- [39] Susan Iversen, Leslie Iversen, and Clifford B Saper. The autonomic nervous system and the hypothalamus. *Principles of Neural Science*, 4:960–981, 2000.
- [40] Yvonne M Ulrich-Lai and James P Herman. Neural regulation of endocrine and autonomic stress responses. *Nature Reviews Neuroscience*, 10(6):397–409, 2009.
- [41] Ian Gibbins. Functional organization of autonomic neural pathways. *Organogenesis*, 9(3):169–175, 2013.
- [42] Naomi I Eisenberger and Steve W Cole. Social neuroscience and health: neurophysiological mechanisms linking social ties with physical health. *Nature Neuroscience*, 15(5): 669–674, 2012.
- [43] E Ron De Kloet, Marian Joëls, and Florian Holsboer. Stress and the brain: from adaptation to disease. *Nature Reviews Neuroscience*, 6(6):463–475, 2005.

- [44] Richard Kvetnansky, Esther L Sabban, and Miklos Palkovits. Catecholaminergic systems in stress: structural and molecular genetic approaches. *Physiological Reviews*, 89(2):535–606, 2009.
- [45] Mark S Gilzenrat, Sander Nieuwenhuis, Marieke Jepma, and Jonathan D Cohen. Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective, & Behavioral Neuroscience*, 10(2):252–269, 2010.
- [46] Marieke Jepma and Sander Nieuwenhuis. Pupil diameter predicts changes in the exploration–exploitation trade-off: Evidence for the adaptive gain theory. *Journal of Cognitive Neuroscience*, 23(7):1587–1596, 2011.
- [47] Jonas Bergquist, Agnieszka Ściubisz, Agnieszka Kaczor, and Jerzy Silberring. Catecholamines and methods for their identification and quantitation in biological tissues and fluids. *Journal of Neuroscience Methods*, 113(1):1–13, 2002.
- [48] Joana Bicker, Ana Fortuna, Gilberto Alves, and Amilcar Falcao. Liquid chromatographic methods for the quantification of catecholamines and their metabolites in several biological samples—a review. *Analytica Chimica Acta*, 768:12–34, 2013.
- [49] Hans Weil-Malherbe, Julius Axelrod, and Robert Tomchick. Blood-brain barrier for adrenaline. *Science*, 129(3357):1226–1227, 1959.
- [50] Steven F Maier and Linda R Watkins. Stressor controllability and learned helplessness: the roles of the dorsal raphe nucleus, serotonin, and corticotropin-releasing factor. *Neuroscience & Biobehavioral Reviews*, 29(4-5):829–841, 2005.
- [51] David A Morilak, Gabe Barrera, David J Echevarria, et al. Role of brain norepinephrine in the behavioral response to stress. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 29(8):1214–1224, 2005.
- [52] Yukiori Goto, Satoru Otani, and Anthony A Grace. The yin and yang of dopamine release: a new perspective. *Neuropharmacology*, 53(5):583–587, 2007.
- [53] Robert Adamec, Andrew Holmes, and Jacqueline Blundell. Vulnerability to lasting anxiogenic effects of brief exposure to predator stimuli: Sex, serotonin and other factors—relevance to ptsd. *Neuroscience & Biobehavioral Reviews*, 32(7):1287–1292, 2008.
- [54] Ellen de Graaf-Roelfsema, Hans A Keizer, Eric van Breda, et al. Hormonal responses to acute exercise, training and overtraining a review with emphasis on the horse. *Veterinary Quarterly*, 29(3):82–101, 2007.
- [55] Jaap M Koolhaas, Alessandro Bartolomucci, Bauke Buwalda, et al. Stress revisited: a critical evaluation of the stress concept. *Neuroscience & Biobehavioral Reviews*, 35(5):1291–1301, 2011.
- [56] James P Herman, Helmer Figueiredo, Nancy K Mueller, et al. Central mechanisms of stress integration: hierarchical circuitry controlling hypothalamo–pituitary–adrenocortical responsiveness. *Frontiers in Neuroendocrinology*, 24(3):151–180, 2003.

- [57] Dimphena CE de Goeij, Richard Kvetnansky, Mark H Whitnall, et al. Repeated stress-induced activation of corticotropin-releasing factor neurons enhances vasopressin stores and colocalization with corticotropin-releasing factor in the median eminence of rats. *Neuroendocrinology*, 53(2):150–159, 1991.
- [58] Andy K Lee, Frederick W Tse, and Amy Tse. Arginine vasopressin potentiates the stimulatory action of crh on pituitary corticotropes via a protein kinase c-dependent reduction of the background trek-1 current. *Endocrinology*, 156(10):3661–3672, 2015.
- [59] Niamh X Cawley, Zhaojin Li, and Y Peng Loh. Biosynthesis, trafficking and secretion of pro-opiomelanocortin-derived peptides. *Journal of Molecular Endocrinology*, 56(4):T77, 2016.
- [60] Robert L Spencer and Terrence Deak. A users guide to hpa axis research. *Physiology & Behavior*, 178:43–65, 2017.
- [61] KV Thrivikraman, Charles B Nemeroff, and Paul M Plotsky. Sensitivity to glucocorticoid-mediated fast-feedback regulation of the hypothalamic–pituitary–adrenal axis is dependent upon stressor specific neurocircuitry. *Brain Research*, 870(1-2):87–101, 2000.
- [62] Chad D Osterlund, Mariana Rodriguez-Santiago, Elizabeth R Woodruff, et al. Glucocorticoid fast feedback inhibition of stress-induced acth secretion in the male rat: rate independence and stress-state resistance. *Endocrinology*, 157(7):2785–2798, 2016.
- [63] Francesca Spiga, Jamie J Walker, John R Terry, and Stafford L Lightman. Hpa axis-rhythms. *Comprehensive Physiology*, 4(3):1273–1298, 2011.
- [64] A Kalsbeek, R Van der Spek, J Lei, et al. Circadian rhythms in the hypothalamo–pituitary–adrenal (hpa) axis. *Molecular and Cellular Endocrinology*, 349(1):20–29, 2012.
- [65] RJ Windle, SA Wood, SL Lightman, and CD Ingram. The pulsatile characteristics of hypothalamo-pituitary-adrenal activity in female lewis and fischer 344 rats and its relationship to differential stress responses. *Endocrinology*, 139(10):4044–4052, 1998.
- [66] Edo Ronald de Kloet and Marian Joëls. Mineralocorticoid receptors and glucocorticoid receptors in hpa stress responses during coping and adaptation. In *Oxford Research Encyclopedia of Neuroscience*. 2020.
- [67] John W Funder. Mineralocorticoid receptors: distribution and activation. *Heart Failure Reviews*, 10:15–22, 2005.
- [68] Matthew DB Claydon and Becky L Conway-Campbell. The glucocorticoid-mediated genomic stress response. *Current Opinion in Endocrine and Metabolic Research*, 25:100363, 2022.
- [69] Huanqing Yang, Sowmya Narayan, and Mathias V Schmidt. From ligands to behavioral outcomes: Understanding the role of mineralocorticoid receptors in brain function. *Stress*, (just-accepted):1–53, 2023.

- [70] Karen Chapman, Megan Holmes, and Jonathan Seckl. 11β -hydroxysteroid dehydrogenases: intracellular gate-keepers of tissue glucocorticoid action. *Physiological Reviews*, 93(3):1139–1206, 2013.
- [71] Maureen E Keller-Wood and Mary F Dallman. Corticosteroid inhibition of acth secretion. *Endocrine Reviews*, 5(1):1–24, 1984.
- [72] Alan G Watts. Glucocorticoid regulation of peptide genes in neuroendocrine crh neurons: a complexity beyond negative feedback. *Frontiers in Neuroendocrinology*, 26(3-4):109–130, 2005.
- [73] Nikolaos P Daskalakis, Onno C Meijer, and E Ron de Kloet. Mineralocorticoid receptor and glucocorticoid receptor work alone and together in cell-type-specific manner: Implications for resilience prediction and targeted therapy. *Neurobiology of Stress*, 18:100455, 2022.
- [74] Christoph M Bamberger, Heinrich M Schulte, and George P Chrousos. Molecular determinants of glucocorticoid receptor function and tissue sensitivity to glucocorticoids. *Endocrine Reviews*, 17(3):245–261, 1996.
- [75] Robert John Denver. Structural and functional evolution of vertebrate neuroendocrine stress systems. *Annals of the New York Academy of Sciences*, 1163(1):1–16, 2009.
- [76] Joyce Cheung and David F Smith. Molecular chaperone interactions with steroid receptors: an update. *Molecular Endocrinology*, 14(7):939–946, 2000.
- [77] Anthony L Fink. Chaperone-mediated protein folding. *Physiological Reviews*, 79(2):425–449, 1999.
- [78] Didier Picard. Chaperoning steroid hormone action. *Trends in Endocrinology & Metabolism*, 17(6):229–235, 2006.
- [79] Jeremy D Baker, Ilayda Ozsan, Santiago Rodriguez Ospina, et al. Hsp90 heterocomplexes regulate steroid hormone receptors: from stress response to psychiatric disease. *International Journal of Molecular Sciences*, 20(1):79, 2018.
- [80] Cordelia Schiene-Fischer and Chao Yu. Receptor accessory folding helper enzymes: the functional role of peptidyl prolyl cis/trans isomerases. *FEBS Letters*, 495(1-2):1–6, 2001.
- [81] Keisuke Maeda, Makoto Habara, Mitsuyasu Kawaguchi, et al. Fkbp51 and fkbp52 regulate androgen receptor dimerization and proliferation in prostate cancer cells. *Molecular Oncology*, 16(4):940–956, 2022.
- [82] William B Pratt, Mario D Galigniana, Jennifer M Harrell, and Donald B DeFranco. Role of hsp90 and the hsp90-binding immunophilins in signalling protein movement. *Cellular Signalling*, 16(8):857–872, 2004.
- [83] Gabriela M Wochnik, Joëlle Ruegg, G Alexander Abel, et al. Fk506-binding proteins 51 and 52 differentially regulate dynein interaction and nuclear translocation of the glucocorticoid receptor in mammalian cells. *Journal of Biological Chemistry*, 280(6):4609–4616, 2005.

- [84] Harry Vermeer, Brenda I Hendriks-Stegeman, Bart van der Burg, et al. Glucocorticoid-induced increase in lymphocytic fkbp51 messenger ribonucleic acid expression: a potential marker for glucocorticoid sensitivity, potency, and bioavailability. *The Journal of Clinical Endocrinology & Metabolism*, 88(1):277–284, 2003.
- [85] Erick T Tatro, Ian P Overall, Marcus Kaul, and Cristian L Achim. Modulation of glucocorticoid receptor nuclear translocation in neurons by immunophilins fkbp51 and fkbp52: implications for major depressive disorder. *Brain Research*, 1286:1–12, 2009.
- [86] Bruce S McEwen and John C Wingfield. The concept of allostasis in biology and biomedicine. *Hormones and Behavior*, 43(1):2–15, 2003.
- [87] Elisabeth B Binder. The role of fkbp5, a co-chaperone of the glucocorticoid receptor in the pathogenesis and therapy of affective and anxiety disorders. *Psychoneuroendocrinology*, 34:S186–S195, 2009.
- [88] Isabella Heuser, Alexander Yassouridis, and Florian Holsboer. The combined dexamethasone/crh test: a refined laboratory test for psychiatric disorders. *Journal of Psychiatric Research*, 28(4):341–356, 1994.
- [89] Bente C Appelhof, Jochanan Huyser, Mijke Verweij, et al. Glucocorticoids and relapse of major depression (dexamethasone/corticotropin-releasing hormone test in relation to relapse of major depression). *Biological Psychiatry*, 59(8):696–701, 2006.
- [90] Patrick F Sullivan, Michael C Neale, and Kenneth S Kendler. Genetic epidemiology of major depression: review and meta-analysis. *American Journal of Psychiatry*, 157(10):1552–1562, 2000.
- [91] Gitta H Lubke, Jouke Jan Hottenga, Raymond Walters, et al. Estimating the genetic variance of major depressive disorder due to all single nucleotide polymorphisms. *Biological Psychiatry*, 72(8):707–709, 2012.
- [92] Elżbieta Gałęcka, Janusz Szemraj, Małgorzata Bieńkiewicz, et al. Single nucleotide polymorphisms of nr3c1 gene and recurrent depressive disorder in population of poland. *Molecular Biology Reports*, 40:1693–1699, 2013.
- [93] MD Klok, EJ Giltay, AJW Van der Does, et al. A common and functional mineralocorticoid receptor haplotype enhances optimism and protects against depression in females. *Translational Psychiatry*, 1(12):e62–e62, 2011.
- [94] Avshalom Caspi, Karen Sugden, Terrie E Moffitt, et al. Influence of life stress on depression: moderation by a polymorphism in the 5-htt gene. *Science*, 301(5631):386–389, 2003.
- [95] Martin A Kohli, Susanne Lucae, Philipp G Saemann, et al. The neuronal transporter gene slc6a15 confers risk to major depression. *Neuron*, 70(2):252–265, 2011.
- [96] Dirk Van West, Jurgen Del-Favero, Y Aulchenko, et al. A major snp haplotype of the arginine vasopressin 1b receptor protects against recurrent major depression. *Molecular Psychiatry*, 9(3):287–292, 2004.

- [97] Elisabeth B Binder, Daria Salyakina, Peter Lichtner, et al. Polymorphisms in *fkbp5* are associated with increased recurrence of depressive episodes and rapid response to antidepressant treatment. *Nature Genetics*, 36(12):1319–1325, 2004.
- [98] T Xie, Maria G Stathopoulou, F De Andrés, et al. Vegf-related polymorphisms identified by gwas and risk for major depression. *Translational Psychiatry*, 7(3):e1055–e1055, 2017.
- [99] Daniel F Levey, Murray B Stein, Frank R Wendt, et al. Bi-ancestral depression gwas in the million veteran program and meta-analysis in 1.2 million individuals highlight new therapeutic directions. *Nature Neuroscience*, 24(7):954–963, 2021.
- [100] Elisabeth B Binder, Rebekah G Bradley, Wei Liu, et al. Association of *fkbp5* polymorphisms and childhood abuse with risk of posttraumatic stress disorder symptoms in adults. *JAMA*, 299(11):1291–1305, 2008.
- [101] Torsten Klengel, Divya Mehta, Christoph Anacker, et al. Allele-specific *fkbp5* dna demethylation mediates gene–childhood trauma interactions. *Nature Neuroscience*, 16(1):33–41, 2013.
- [102] Charlotte Elisabeth Piechaczek, Ellen Greimel, Lisa Feldmann, et al. Interactions between *fkbp5* variation and environmental stressors in adolescent major depression. *Psychoneuroendocrinology*, 106:28–37, 2019.
- [103] Jee In Kang, Tae Yong Kim, Jin Hee Choi, et al. Allele-specific dna methylation level of *fkbp5* is associated with post-traumatic stress disorder. *Psychoneuroendocrinology*, 103:1–7, 2019.
- [104] Johanna Klinger-König, Johannes Hertel, Sandra Van der Auwera, et al. Methylation of the *fkbp5* gene in association with *fkbp5* genotypes, childhood maltreatment and depression. *Neuropsychopharmacology*, 44(5):930–938, 2019.
- [105] Zhengxiao Fan, Hong Zhu, Tingting Zhou, et al. Using the tube test to measure social hierarchy in mice. *Nature Protocols*, 14(3):819–831, 2019.
- [106] Jacqueline S Womersley, Simone Roeh, Lindi Martin, et al. *fkbp5* intron 7 methylation is associated with higher anxiety proneness and smaller right thalamus volume in adolescents. *Brain Structure and Function*, 227(8):2809–2820, 2022.
- [107] Alex Ferrer, Javier Costas, Javier Labad, et al. *Fkbp5* polymorphisms and hypothalamic-pituitary-adrenal axis negative feedback in major depression and obsessive-compulsive disorder. *Journal of Psychiatric Research*, 104:227–234, 2018.
- [108] Niat T Gebru, Shannon E Hill, and Laura J Blair. Genetically engineered mouse models of *fk506*-binding protein 5. *Journal of Cellular Biochemistry*, 2023.
- [109] Sebastian H Scharf, Claudia Liebl, Elisabeth B Binder, et al. Expression and regulation of the *fkbp5* gene in the adult mouse brain. *PLOS ONE*, 6(2):e16883, 2011.
- [110] *Allen Mouse Brain Atlas* [[http:// mouse.brain-map.org/](http://mouse.brain-map.org/)]. Institute for Brain Science. 2011.

- [111] Gianluigi Guidotti, Francesca Calabrese, Christoph Anacker, et al. Glucocorticoid receptor and fkbp5 expression is altered following exposure to chronic stress: modulation by antidepressant treatment. *Neuropsychopharmacology*, 38(4):616–627, 2013.
- [112] Sara Santarelli, Christoph Zimmermann, Georgia Kalideris, et al. An adverse early life environment can enhance stress resilience in adulthood. *Psychoneuroendocrinology*, 78: 213–221, 2017.
- [113] Jonathan J Sabbagh, John C O’Leary III, Laura J Blair, et al. Age-associated epigenetic upregulation of the fkbp5 gene selectively impairs stress resiliency. *PLOS ONE*, 9(9): e107241, 2014.
- [114] Georgia Balsevich, Alexander S Häusl, Carola W Meyer, et al. Stress-responsive fkbp51 regulates akt2-as160 signaling and metabolic function. *Nature Communications*, 8(1): 1725, 2017.
- [115] Chadi Touma, Nils Christian Gassen, Leonie Herrmann, et al. Fk506 binding protein 5 shapes stress responsiveness: modulation of neuroendocrine reactivity and coping behavior. *Biological Psychiatry*, 70(10):928–936, 2011.
- [116] John C O’Leary III, Sheetal Dharia, Laura J Blair, et al. A new anti-depressive strategy for the elderly: ablation of fkbp5/fkbp51. *PloS one*, 6(9):e24840, 2011.
- [117] Lianne Hoeijmakers, Daniela Harbich, Bianca Schmid, et al. Depletion of fkbp51 in female mice shapes hpa axis activity. *PLOS ONE*, 9(4):e95796, 2014.
- [118] Stefana Albu, Christoph PN Romanowski, M Letizia Curzi, et al. Deficiency of fk 506-binding protein (fkbp) 51 alters sleep architecture and recovery sleep responses to stress in mice. *Journal of Sleep Research*, 23(2):176–185, 2014.
- [119] Jakob Hartmann, Klaus V. Wagner, Claudia Liebl, et al. The involvement of fk506-binding protein 51 (fkbp5) in the behavioral and neuroendocrine effects of chronic social defeat stress. *Neuropharmacology*, 62:332–339, 2012.
- [120] Benjamin K Attwood, Julie-Myrtille Bourgognon, Satyam Patel, et al. Neuropsin cleaves ephb2 in the amygdala to control anxiety. *Nature*, 473(7347):372–375, 2011.
- [121] Rita J Valentino, Stephen L Foote, and Gary Aston-Jones. Corticotropin-releasing factor activates noradrenergic neurons of the locus coeruleus. *Brain Research*, 270(2): 363–367, 1983.
- [122] LW Swanson and BK Hartman. The central adrenergic system. an immunofluorescence study of the location of cell bodies and their efferent connections in the rat utilizing dopamine-b-hydroxylase as a marker. *Journal of Comparative Neurology*, 163(4):467–505, 1975.
- [123] JC Reil. Untersuchungen über den bau des grossen gehirns im menschen. *Arch Physiol*, 9:136–208, 1809.
- [124] Joseph Wenzel and Karl Wenzel. *De penitiori structura cerebri hominis et brutorum*, volume 1. Cotta, 1812.

- [125] Dan J Chandler, Patricia Jensen, Jordan G McCall, et al. Redefining noradrenergic neuromodulation of behavior: impacts of a modular locus coeruleus architecture. *Journal of Neuroscience*, 39(42):8239–8249, 2019.
- [126] KL Double, VN Dedov, H Fedorow, et al. The comparative biology of neuromelanin and lipofuscin in the human brain. *Cellular and Molecular Life Sciences*, 65:1669–1682, 2008.
- [127] Toshiharu Nagatsu, Akira Nakashima, Hirohisa Watanabe, et al. The role of tyrosine hydroxylase as a key player in neuromelanin synthesis and the association of neuromelanin with parkinson’s disease. *Journal of Neural Transmission*, pages 1–15, 2023.
- [128] Yukti Sharma, Tao Xu, Werner M Graf, et al. Comparative anatomy of the locus coeruleus in humans and nonhuman primates. *Journal of Comparative Neurology*, 518(7):963–971, 2010.
- [129] Stephen L Foote and John H Morrison. Extrathalamic modulation of cortical function. *Annual Review of Neuroscience*, 10(1):67–95, 1987.
- [130] Lindsay A Schwarz, Kazunari Miyamichi, Xiaojing J Gao, et al. Viral-genetic tracing of the input–output organization of a central noradrenaline circuit. *Nature*, 524(7563):88–92, 2015.
- [131] AFT Arnsten and PS Goldman-Rakic. Selective prefrontal cortical projections to the region of the locus coeruleus and raphe nuclei in the rhesus monkey. *Brain Research*, 306(1-2):9–18, 1984.
- [132] Jodi L Downs, Michael R Dunn, Erzsebet Borok, et al. Orexin neuronal changes in the locus coeruleus of the aging rhesus macaque. *Neurobiology of Aging*, 28(8):1286–1295, 2007.
- [133] Rita J Valentino and Elisabeth Van Bockstaele. Convergent regulation of locus coeruleus activity as an adaptive response to stress. *European Journal of Pharmacology*, 583(2-3):194–203, 2008.
- [134] Jordan G McCall, Ream Al-Hasani, Edward R Siuda, et al. Crh engagement of the locus coeruleus noradrenergic system mediates stress-induced anxiety. *Neuron*, 87(3):605–620, 2015.
- [135] Myung-A Kim, Hyun S Lee, Beob Y Lee, and Barry D Waterhouse. Reciprocal connections between subdivisions of the dorsal raphe and the nuclear core of the locus coeruleus in the rat. *Brain Research*, 1026(1):56–67, 2004.
- [136] Susan J Sara. The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews Neuroscience*, 10(3):211–223, 2009.
- [137] SL Foote, G Aston-Jones, and FE Bloom. Impulse activity of locus coeruleus neurons in awake rats and monkeys is a function of sensory stimulation and arousal. *Proceedings of the National Academy of Sciences*, 77(5):3033–3037, 1980.

- [138] G Aston-Jones and FE6564235 Bloom. Activity of norepinephrine-containing locus coeruleus neurons in behaving rats anticipates fluctuations in the sleep-waking cycle. *Journal of Neuroscience*, 1(8):876–886, 1981.
- [139] Gary Aston-Jones, Janusz Rajkowski, Piotr Kubiak, and Tatiana Alexinsky. Locus coeruleus neurons in monkey are selectively activated by attended cues in a vigilance task. *Journal of Neuroscience*, 14(7):4467–4480, 1994.
- [140] Barry D Waterhouse and Donald J Woodward. Interaction of norepinephrine with cerebrocortical activity evoked by stimulation of somatosensory afferent pathways in the rat. *Experimental Neurology*, 67(1):11–34, 1980.
- [141] Jim McBurney-Lin, Ju Lu, Yi Zuo, and Hongdian Yang. Locus coeruleus-norepinephrine modulation of sensory processing and perception: A focused review. *Neuroscience & Biobehavioral Reviews*, 105:190–199, 2019.
- [142] Stefan Hirschberg, Yong Li, Andrew Randall, et al. Functional dichotomy in spinal-vs prefrontal-projecting locus coeruleus modules splits descending noradrenergic analgesia from ascending aversion and anxiety in rats. *Elife*, 6:e29808, 2017.
- [143] Devin Mueller, James T Porter, and Gregory J Quirk. Noradrenergic signaling in infralimbic cortex increases cell excitability and strengthens memory for fear extinction. *Journal of Neuroscience*, 28(2):369–375, 2008.
- [144] Akira Uematsu, Bao Zhen Tan, Edgar A Ycu, et al. Modular organization of the brainstem noradrenaline system coordinates opposing learning states. *Nature Neuroscience*, 20(11):1602–1611, 2017.
- [145] Stephen T Mason and Hans C Fibiger. Regional topography within noradrenergic locus coeruleus as revealed by retrograde transport of horseradish peroxidase. *Journal of Comparative Neurology*, 187(4):703–724, 1979.
- [146] SE Loughlin, SL Foote, and FE Bloom. Efferent projections of nucleus locus coeruleus: topographic organization of cells of origin demonstrated by three-dimensional reconstruction. *Neuroscience*, 18(2):291–306, 1986.
- [147] Eduardo E Benarroch. Locus coeruleus. *Cell and Tissue Research*, 373(1):221–232, 2018.
- [148] Gina R Poe, Stephen Foote, Oxana Eschenko, et al. Locus coeruleus: a new look at the blue spot. *Nature Reviews Neuroscience*, 21(11):644–659, 2020.
- [149] Sabrina D Robertson, Nicholas W Plummer, and Patricia Jensen. Uncovering diversity in the development of central noradrenergic neurons and their efferents. *Brain Research*, 1641:234–244, 2016.
- [150] Gary Aston-Jones and Jonathan D Cohen. An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annual Review Neuroscience*, 28:403–450, 2005.

- [151] David M Devilbiss and Barry D Waterhouse. Phasic and tonic patterns of locus coeruleus output differentially modulate sensory network function in the awake rat. *Journal of Neurophysiology*, 105(1):69–87, 2011.
- [152] Craig W Berridge and Barry D Waterhouse. The locus coeruleus–noradrenergic system: modulation of behavioral state and state-dependent cognitive processes. *Brain Research Reviews*, 42(1):33–84, 2003.
- [153] Rita J Valentino, Michelle E Page, and Andre L Curtis. Activation of noradrenergic locus coeruleus neurons by hemodynamic stress is due to local release of corticotropin-releasing factor. *Brain Research*, 555(1):25–34, 1991.
- [154] Erno J Hermans, Hein JF Van Marle, Lindsey Ossewaarde, et al. Stress-related noradrenergic activity prompts large-scale neural network reconfiguration. *Science*, 334(6059):1151–1153, 2011.
- [155] Thomas F Giustino, Karthik R Ramanathan, Michael S Totty, et al. Locus coeruleus norepinephrine drives stress-induced increases in basolateral amygdala firing and impairs extinction learning. *Journal of Neuroscience*, 40(4):907–916, 2020.
- [156] Jordan G McCall, Edward R Siuda, Dionnet L Bhatti, et al. Locus coeruleus to basolateral amygdala noradrenergic projections promote anxiety-like behavior. *Elife*, 6:e18247, 2017.
- [157] AL Curtis, LA Pavcovich, DE Grigoriadis, and RJ Valentino. Previous stress alters corticotropin-releasing factor neurotransmission in the locus coeruleus. *Neuroscience*, 65(2):541–550, 1995.
- [158] Andre L Curtis, Luis A Pavcovich, and Rita J Valentino. Long-term regulation of locus coeruleus sensitivity to corticotropin-releasing factor by swim stress. *Journal of Pharmacology and Experimental Therapeutics*, 289(3):1211–1219, 1999.
- [159] Jay M Weiss, Julie C Stout, Mark F Aaron, et al. Depression and anxiety: role of the locus coeruleus and corticotropin-releasing factor. *Brain Research Bulletin*, 35(5-6):561–572, 1994.
- [160] Jaanus Harro and Lars Oreland. Depression as a spreading adjustment disorder of monoaminergic neurons: a case for primary implication of the locus coeruleus. *Brain Research Reviews*, 38(1-2):79–128, 2001.
- [161] Laurel S Morris, Jordan G McCall, Dennis S Charney, and James W Murrough. The role of the locus coeruleus in the generation of pathological anxiety. *Brain and Neuroscience Advances*, 4:2398212820930321, 2020.
- [162] Alexander Mathis, Pranav Mamidanna, Kevin M Cury, et al. Deeplabcut: markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, 21(9):1281–1289, 2018.
- [163] Jessy Lauer, Mu Zhou, Shaokai Ye, et al. Multi-animal pose estimation, identification and tracking with deeplabcut. *Nature Methods*, 19(4):496–504, 2022.

- [164] Jacob M Graving, Daniel Chae, Hemal Naik, et al. Deepposekit, a software toolkit for fast and robust animal pose estimation using deep learning. *Elife*, 8:e47994, 2019.
- [165] Talmo D Pereira, Diego E Aldarondo, Lindsay Willmore, et al. Fast animal pose estimation using deep neural networks. *Nature Methods*, 16(1):117–125, 2019.
- [166] Talmo D Pereira, Nathaniel Tabris, Arie Matsliah, et al. Slep: A deep learning system for multi-animal pose tracking. *Nature Methods*, 19(4):486–495, 2022.
- [167] Markus Marks, Qiuhan Jin, Oliver Sturman, et al. Deep-learning-based identification, tracking, pose estimation and behaviour classification of interacting primates and mice in complex environments. *Nature Machine Intelligence*, 4(4):331–340, 2022.
- [168] Simon RO Nilsson, Nastacia L Goodwin, Jia Jie Choong, et al. Simple behavioral analysis (simba)—an open source toolkit for computer classification of complex social behaviors in experimental animals. *BioRxiv*, pages 2020–04, 2020.
- [169] Cristina Segalin, Jalani Williams, Tomomi Karigo, et al. The mouse action recognition system (mars) software pipeline for automated analysis of social behaviors in mice. *Elife*, 10:e63720, 2021.
- [170] Joeri Bordes, Lucas Miranda, Maya Reinhardt, et al. Automatically annotated motion tracking identifies a distinct social behavioral profile following chronic social defeat stress. *Nature Communications*, 14(1):4319, 2023.
- [171] Alexander I Hsu and Eric A Yttri. B-soid, an open-source unsupervised algorithm for identification and fast prediction of behaviors. *Nature Communications*, 12(1):5188, 2021.
- [172] Eric Yttri, Alexander Hsu, Jens Schweihoff, and Martin Schwarz. A-soid, an active learning platform for expert-guided, data efficient discovery of behavior. 2023.
- [173] Alexander B Wiltschko, Tatsuya Tsukahara, Ayman Zeine, et al. Revealing the structure of pharmacobehavioral space through motion sequencing. *Nature Neuroscience*, 23(11):1433–1443, 2020.
- [174] Caleb Weinreb, Mohammed Abdal Monium Osman, Libby Zhang, et al. Keypoint-moseq: parsing behavior by linking point tracking to pose dynamics. *BioRxiv*, pages 2023–03, 2023.
- [175] Kevin Luxem, Petra Mocellin, Falko Fuhrmann, et al. Identifying behavioral structure from deep variational embeddings of animal motion. *Communications Biology*, 5(1):1267, 2022.
- [176] Thomas J Upton, Eder Zavala, Paal Methlie, et al. High-resolution daily profiles of tissue adrenal steroids by portable automated collection. *Science Translational Medicine*, 15(701):eadg8464, 2023.
- [177] Elmira Anderzhanova, Kathrin Hafner, Andreas J Genewsky, et al. The stress susceptibility factor fkbp51 controls s-ketamine-evoked release of mbdnf in the prefrontal cortex of mice. *Neurobiology of Stress*, 13:100239, 2020.

- [178] Jiesi Feng, Changmei Zhang, Julieta E Lischinsky, et al. A genetically encoded fluorescent sensor for rapid and specific in vivo detection of norepinephrine. *Neuron*, 102(4):745–761, 2019.
- [179] James J Jun, Nicholas A Steinmetz, Joshua H Siegle, et al. Fully integrated silicon probes for high-density recording of neural activity. *Nature*, 551(7679):232–236, 2017.
- [180] Thomas Zhihao Luo, Adrian Gopnik Bondy, Diksha Gupta, et al. An approach for long-term, multi-probe neuropixels recordings in unrestrained rats. *Elife*, 9:e59716, 2020.
- [181] Séverine Durand, Gregory R Heller, Tamina K Ramirez, et al. Acute head-fixed recordings in awake mice with multiple neuropixels probes. *Nature Protocols*, 18(2):424–457, 2023.
- [182] James A Hope, Travis Beckerle, Pin-Hao Cheng, et al. Brain-wide neural recordings in mice navigating physical spaces enabled by a cranial exoskeleton. *BioRxiv*, pages 2023–06, 2023.
- [183] Steffen Schneider, Jin Hwa Lee, and Mackenzie Weygandt Mathis. Learnable latent embeddings for joint behavioural and neural analysis. *Nature*, pages 1–9, 2023.
- [184] Juan Pablo Lopez, Elena Brivio, Alice Santambrogio, et al. Single-cell molecular profiling of all three components of the hpa axis reveals adrenal abcb1 as a regulator of stress adaptation. *Science Advances*, 7(5):eabe4497, 2021.
- [185] Clara Engelhardt, Fiona Tang, Radwa Elkhateib, et al. Fkbp51 in the oval bed nucleus of the stria terminalis regulates anxiety-like behavior. *Eneuro*, 8(6), 2021.
- [186] Lotte van Doeselaar, Tibor Stark, Shiladitya Mitra, et al. Sex-specific and opposed effects of fkbp51 in glutamatergic and gabaergic neurons: Implications for stress susceptibility and resilience. *Proceedings of the National Academy of Sciences*, 120(23):e2300722120, 2023.
- [187] Robert J Handa, Loyd H Burgess, Janice E Kerr, and Joan A O’Keefe. Gonadal steroid hormone receptors and sex differences in the hypothalamo-pituitary-adrenal axis. *Hormones and behavior*, 28(4):464–476, 1994.
- [188] Debra A Bangasser, Andre Curtis, Beverly AS Reyes, et al. Sex differences in corticotropin-releasing factor receptor signaling and trafficking: potential role in female vulnerability to stress-related psychopathology. *Molecular Psychiatry*, 15(9):896–904, 2010.
- [189] Debra A Bangasser, Kimberly R Wiersielis, and Sabina Khantsis. Sex differences in the locus coeruleus-norepinephrine system and its regulation by stress. *Brain Research*, 1641:177–188, 2016.
- [190] Rachel-Karson Thériault, Joshua D Manduca, and Melissa L Perreault. Sex differences in innate and adaptive neural oscillatory patterns link resilience and susceptibility to chronic stress in rats. *Journal of Psychiatry and Neuroscience*, 46(2):E258–E270, 2021.

- [191] Harrison L Blount, Jason Dee, Lizhen Wu, et al. Resilience to anxiety and anhedonia after predator scent stress is accompanied by increased nucleus accumbens mglu5 in female rats. *BioRxiv*, pages 2022–04, 2022.
- [192] Dongyu Wei, Vaishali Talwar, and Dayu Lin. Neural circuits of social behaviors: innate yet flexible. *Neuron*, 109(10):1600–1620, 2021.
- [193] Alexander Z Harris, Piray Atsak, Zachary H Bretton, et al. A novel method for chronic social defeat stress in female mice. *Neuropsychopharmacology*, 43(6):1276–1283, 2018.
- [194] Lotte van Doeselaar, Huanqing Yang, Joeri Bordes, et al. Chronic social defeat stress in female mice leads to sex-specific behavioral and neuroendocrine effects. *Stress*, 24(2):168–180, 2021.
- [195] Andrea Harée Pantoja-Urbán, Samuel Richer, Amelie Mittermaier, et al. Gains and losses: Resilience to social defeat stress in adolescent female mice. *Biological Psychiatry*, 2023.
- [196] Sam A Golden, Herbert E Covington III, Olivier Berton, and Scott J Russo. A standardized protocol for repeated social defeat stress in mice. *Nature Protocols*, 6(8):1183–1191, 2011.
- [197] Koichi Hashikawa, Yoshiko Hashikawa, Julieta Lischinsky, and Dayu Lin. The neural mechanisms of sexually dimorphic aggressive behaviors. *Trends in Genetics*, 34(10):755–776, 2018.
- [198] Vaishnav Krishnan, Ming-Hu Han, Danielle L Graham, et al. Molecular adaptations underlying susceptibility and resistance to social defeat in brain reward regions. *Cell*, 131(2):391–404, 2007.
- [199] MV Schmidt, SH Scharf, C Liebl, et al. A novel chronic social stress paradigm in female mice. *Hormones and Behavior*, 57(4-5):415–420, 2010.
- [200] Amber Koert, Annemie Ploeger, Claudi LH Bockting, et al. The social instability stress paradigm in rat and mouse: A systematic review of protocols, limitations, and recommendations. *Neurobiology of Stress*, 15:100410, 2021.
- [201] Joëlle Lopez and Rosemary C Bagot. Defining valid chronic stress models for depression with female rodents. *Biological Psychiatry*, 90(4):226–235, 2021.
- [202] Yair Shemesh, Yehezkel Sztainberg, Oren Forkosh, et al. High-order social interactions in groups of mice. *Elife*, 2:e00759, 2013.
- [203] Fabrice De Chaumont, Elodie Ey, Nicolas Torquet, et al. Real-time analysis of the behaviour of groups of mice via a depth-sensing camera and machine learning. *Nature Biomedical Engineering*, 3(11):930–942, 2019.
- [204] Dean G Kilpatrick, Heidi S Resnick, Melissa E Milanak, et al. National estimates of exposure to traumatic events and ptsd prevalence using dsm-iv and dsm-5 criteria. *Journal of Traumatic Stress*, 26(5):537–547, 2013.

Acknowledgements

The thesis is just a small part of an important period in my scientific and personal life. I am extremely thankful to have been able to share this journey with a lot of special people who have all been in their own way, a crucial part of it.

Mathias, as a mentor and friend you have always provided me with great scientific and personal advice. I started in your lab more than five years ago and have gained a lot important skills and life lessons because of you. The environment and atmosphere that you have created in the lab is something amazing and unique. You are an excellent example as a scientist and I am extremely grateful and proud to have been part of your lab.

Lucas, it has been an absolute pleasure working with you so closely on so many different projects. We understand each other very well and I cherished the way we worked together and let each other do the things we were good at, which I think turned out amazingly. In recent years we have also grown together as good friends and I am very grateful to be able to have such a colleague that I can call my friend. I am so sure that you will achieve great things in your next jobs and I am hoping that we will find ways to still collaborate together, as working together has been a pleasure, always.

Pablo, also you I have been able to call a good friend since recent years. I really enjoyed doing research projects with you. Something I found particular impressive is your transformation from a postdoc mindset to a research group leader within just a couple of months. I am grateful for the faith you have given me and the amazing advice especially regarding my scientific career.

Aron, you are a great scientist and I have enjoyed working together with you. You have also given me great advice regarding my research projects and I am grateful how open and transparent you were on your research projects.

Maya and Larissa, both of you fitted so well in the lab and with me personally. I think we had a great connection, which led to a great output of science. In both of your projects, there were of course moments that felt like the project would never work out (welcome to science), but to me both of you had the correct attitude and mindset to adapt and turn around the

project into something meaningful.

Lea and Lotte, it is so funny to look back at our time now, each with our own flaws and insecurities. Starting our PhDs at the same time can and has given us extra challenges and even fights. But no matter what in the end we were always there to help each other during experiments or during difficult PhD times. I still cannot believe how much we have all learned in the last years and I have us all grow together and I am extremely excited to see what you guys will achieve next!

Max, Alex, Karla and Clara, I am grateful for you guys to welcome me as a master's student in the lab. It was not always easy to be the only non-german colleague in the lab, but the way you all welcomed me has been the reason for me to stay in the lab as a PhD student. I remember how Alex invited me in my first week in the lab to come to his house and join him and his friends to the oktoberfest, but also Max of course who has invited me plenty of times to the Pils doktor and we've had so many great moments! Still waiting to go the champions league final though, hopefully we will make it someday ;)

Sowmya, Veronika, Shila, Huanqing, Margherita, Ben and London, the gang that joined the lab after me, I am happy to see how each of you fitted very within the lab. I am sure you will keep the same spirit going in the lab as before and I am wishing you all good luck with your projects!

Nils, Thomas, Elmira, Jakob, Emily, Simon and Jan, you have all been important parts in my scientific work, thank you so much for you help on my research projects, it would not have been possible without you!

Niek, Sylvie and Janssen, it's awesome to have met you guys at the different conferences over the years. I think those times really give me a boost, first of all to see that you're not the only one struggling, and second to get inspired and how to tackle your scientific problems, and of course to blow of some steam ;)

Wim, Steef, Bas and Floris, you guys have been my friends since beginning of my time here in Munich. For some literally, Floris, remember how we met on one of our first day in Munich at marienplatz, trying to find each other without knowing who the other was? All of you have been such an important part of my life in Munich and also the reason for me to be able to stay and enjoy my time here. Thank you all for that. Steef and Floris, of course you guys already left Munich before and I am so happy how we still try to meet up often. Bas

and Wim, I am going to miss you guys incredibly while leaving Munich, hope we can also be able to keep on seeing each other in a different context.

Lucas, Bas and Tjarda, you guys have been there for me my whole life and having such a strong and great relationship has allowed me to achieve this thesis. Thank you for all your support.

Claire, being together makes me extremely happy and I am so happy that we've been able stay so easily together, even though we've not been able to be in the same place for the majority of our time. I am looking forward to next years with you, thank you for all your support.

Publications

- 2023 **Bordes, J.***, Miranda, L.*, Reinhardt, M., Narayan, S., Hartmann, J., Newman, E.L., Brix, L.M., van Doeselaar, L., Engelhardt, C., Dillmann, L., Mitra, S., Ressler, K.J., Pütz, B., Agakov, F., Müller-Myhsok, B., Schmidt, M.V. **Automatically annotated motion tracking identifies a distinct social behavioral profile following chronic social defeat stress**
Nature Communications
* Shared authorship
- Miranda, L., **Bordes, J.**, Pütz, B., Schmidt, M.V., Müller-Myhsok, B. **DeepOF: a python package for supervised and unsupervised pattern recognition in mice motion tracking data**
The Journal of Open Source Software, 2023, 8(86), 5394
- Bordes, J.**, Miranda, L., Müller-Myhsok, B., Schmidt, M.V. **Advancing social behavioral neuroscience by integrating ethology and comparative psychology methods through machine learning**
Neuroscience and Biobehavioral Reviews, 2023, 151, 105243
- Miranda, L.* , **Bordes, J.***, Gasperoni, S., Lopez, J.P. **Increasing resolution in stress neurobiology: from single cells to complex group behaviors**
Stress, 2023, 26(1), 2186141
*Shared authorship
- van Doeselaar, L., Stark, T., Mitra, S., Yang, H., **Bordes, J.**, Stolwijk, L., Engelhardt, C., Kovarova, V., Narayan, S., Brix, L.M., Springer, M., Deussing, J., Lopez, J.P., Czisch, M., Schmidt, M.V. **Sex-specific and opposed effects of FKBP51 in glutamatergic and GABAergic neurons: Implications for stress susceptibility and resilience**
Proceedings of the National Academy of Sciences, 120(23), e2300722120
- Brix, L.M., Monleon, D., Collado, M.C., Ederveen, T.H.A., Toksöz, I., **Bordes, J.**, van Doeselaar, L., Engelhardt, C., Mitra, S., Narayan, S., Schmidt, M.V. **Metabolic effects of early life stress and pre-pregnancy obesity are longlasting and sex-specific in mice**
European Journal of Neuroscience

- 2022 von Mücke-Heim, I.A., Urbina-Treviño, L., **Bordes, J.**, Ries, C., Schmidt, M.V., Deussing, J. **Introducing a depression-like syndrome for translational neuropsychiatry: a plea for taxonomical validity and improved comparability between humans and mice**
Molecular Psychiatry, 28(1), 329-340.
- Brix, L.M., Toksöz, I., Aman, L., Kovarova, V., Springer, M., **Bordes, J.**, van Doeselaar, L., Engelhardt, C., Häusl, A.S., Narayan, S., Sterlemann, V., Yang, H., Deussing, J.M., Schmidt, M.V. **Contribution of the co-chaperone FKBP51 in the ventromedial hypothalamus to metabolic homeostasis in male and female mice**
Molecular metabolism, 65, 101579.
- Brix, L.M., Häusl, A.S., Toksöz, I., **Bordes, J.**, van Doeselaar, L., Engelhardt, C., Narayan, S., Springer, M., Sterlemann, V., Deussing, J.M., Chen, A., Schmidt, M.V. **The co-chaperone FKBP51 modulates HPA axis activity and age-related maladaptation of the stress system in pituitary proopiomelanocortin cells**
Psychoneuroendocrinology, 138, 105670.
- 2021 Engelhardt, C., Tang, F., Elkhateib, R., **Bordes, J.**, Brix, L.M., van Doeselaar, L., Häusl, A.S., Pöhlmann, M.L., Schraut, K.G., Yang, H., Chen, A., Deussing, J., Schmidt, M.V. **FKBP51 in the Oval Bed Nucleus of the Stria Terminalis Regulates Anxiety-Like Behavior**
eNeuro, 8 (6).
- 2020 van Doeselaar, L., Yang, H., **Bordes, J.**, Brix, L.M., Engelhardt, C., Tang, F., Schmidt, M. V. **Chronic social defeat stress in female mice leads to sex-specific behavioral and neuroendocrine effects.**
Stress, 24(2), 168-180.

Assertion/Eidesstattliche Erklärung

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbständig und ohne unerlaubte Hilfe angefertigt ist.

Joeri Bordes
München, den 05.07.2023

Erklärung

Hiermit erkläre ich,
dass die Dissertation nicht ganz oder in wesentlichen Teilen einer anderen Prüfungskommission vorgelegt worden ist.

dass ich mich anderweitig einer Doktorprüfung ohne Erfolg nicht unterzogen habe.

Joeri Bordes
München, den 05.07.2023

