Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Target-biased DP output predicts divergent social strategies after witnessed stress

Huilin Hong

honghuilin@fudan.edu.cn

Huadong Hospital

Wei Ren

Shanghai Sixth People's Hospital

Lei Li

Huadong Hospital

Zhiming Li

Huadong Hospital

Jun Zhang

Huadong Hospital

Weisheng Wang

Huadong Hospital

Article

Keywords: vicarious social stress, dorsal peduncular cortex, bed nucleus of the stria terminalis, dorsomedial hypothalamus

Posted Date: November 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11633918/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Target-biased DP output predicts divergent social strategies after witnessed stress

Huilin Hong[1,3*], Wei Ren[2,3], Lei Li[1], Zhiming Li[1], Jun Zhang[1*], Weisheng Wang[1*]

[1]Department of Radiology, Huadong Hospital, Institute for Translational Brain Research, Center for Clinical Neuro-AI, Shanghai Key Laboratory of Clinical Geriatric Medicine, Shanghai Institute of Geriatrics and Gerontology, State Key Laboratory of Brain Function and Disorders, and MOE Frontiers Center for Brain Science, Fudan University, Shanghai, China

[2]Shanghai Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China

[3]These authors contributed equally to this work

*Corresponding Author. Email: Jun Zhang, zhangjun_zj@fudan.edu.cn; Weisheng Wang, weishengw@fudan.edu.cn; Huilin Hong, honghuilin@fudan.edu.cn

Keywords: vicarious social stress; dorsal peduncular cortex; bed nucleus of the stria terminalis; dorsomedial hypothalamus

Abstract

Adverse social experiences often lead to divergent behavioral outcomes, yet the neural mechanisms by which a common stressor drives individual-specific social strategies remain poorly understood. Here, using a witnessed social stress model, we reveal that a shared population of neurons in the dorsal peduncular cortex (DP) governs the emergence of distinct aggression-prone and avoidance-prone states. We found that while DP neurons projecting to the bed nucleus of the stria terminalis (BNST) and dorsomedial hypothalamus (DMH) largely overlap, their activity develops a target-specific output bias during stress observation, prospectively predicting the animal's future social strategy. Crucially, pathway-specific manipulation of DP-BNST or DP-DMH projections during stress causally shifts state probability, and established behavioral biases can be reversed through cell-type-specific rebalancing of these downstream outputs. Our findings establish a mechanistic framework where a shared prefrontal circuit dynamically allocates defensive strategies via relative output bias, highlighting an experience-dependent and modifiable neural basis for individual differences in stress coping.

Introduction

Adverse social experiences are a pervasive feature of life and can produce persistent changes in subsequent behavior (1-3). A central challenge is to understand why a common stressful event is followed by aggression in some individuals and avoidance in others. Recent work shows that aggression can be transmitted through observation and that familiarity with the observed aggressor can strongly gate this effect, while direct social defeat and dominance history also generate heterogeneous later aggression and avoidance (4,21-25,36-41). From an ecological perspective, attack and avoidance are state-dependent strategies whose value depends on controllability, risk, resources and social context (5-7). Neural circuits must therefore adjust defensive weighting as the expected costs and benefits of a social threat change. The computations that establish this balance remain incompletely understood.

The dorsal peduncular cortex (DP) form a contiguous ventromedial frontal territory with projections to hypothalamic and extended-amygdala structures involved in stress physiology and defensive behavior (8-11). It is strategically connected to key hypothalamic and extended-amygdala systems known to be involved in stress physiology and behavioral responses. While previous research has implicated DP-dorsomedial hypothalamus (DMH) signaling in psychosocial stress responses (10) and highlighted the stress-sensitive nature of the bed nucleus of the stria terminalis (BNST) circuits (12,13,34,35), these studies have predominantly focused on average effects or investigated single pathways in isolation. Consequently, it remains unresolved whether the DP implements divergent social strategies through entirely segregated neuronal pathways, or through a shared population whose relative output to the BNST and DMH dynamically determines the subsequent behavioral strategy.

To address this critical gap, we employed a multi-faceted approach combining data-driven behavioral phenotyping, longitudinal calcium imaging, projection tracing, dual-target axonal recording, pathway-specific perturbation, and cell-type-defined rebalancing. Our study was specifically designed to investigate whether witnessed social stress generates a shared DP state, how its relative output to BNST and DMH determines the subsequent social strategy, precisely when this output bias forms, and whether it can be therapeutically rebalanced. By distinguishing experience-dependent state formation from pre-existing behavioral differences, our work provides a mechanistic framework for understanding individual variability in response to social adversity.

Methods and Materials

Animals and study design

Experiments used 7-week-old C57BL/6J Gpt female observer and victim mice and a separate resident-aggressor cohort. Retired male CD-1 breeders (approximately 6 months; Charles River Laboratories) were singly housed. GABA-Cre, vGlut2-Cre and TRAP2;Ai9 lines were used in the indicated experiments. Female mice were housed six per cage under a 12-h light/dark cycle (lights on at 08:00) with food and water available ad libitum. All procedures were approved by Fudan University and complied with applicable national guidelines.

Witnessed social stress and behavioral testing

A screened resident aggressor attacked a victim during a 15-min observation period while the observer remained behind a transparent perforated barrier that prevented direct physical contact but permitted visual, auditory and olfactory cues (Fig. 1A). Observers are familiar with victim and partner animals. Observers completed a 10-min baseline free interaction with a nonaggressive conspecific before observation and a second 10-min interaction afterward (Fig. 1B). Partner mouse are different during pre- and post-interaction.

Resident aggressors received approximately 500 μg testosterone propionate subcutaneously 24 h before behavioral screening. Attack number was monitored across screening sessions (Fig. S2A). For each observation session, attack count, cumulative attack duration and the timing of each attack were recorded. Heart rate was recorded continuously during witnessed stress and synchronized to video. Attack epochs were marked for event-aligned analysis (Fig. 1C), and mouse-level heart-rate measures were compared across later states (Fig. S2C). Baseline and post-stress social behavior was measured during 10-min free interaction with a familiar nonaggressive conspecific. Social behavior including sniffing (defined as nose contact with the demonstrator), allogrooming (defined as visible licking or mouth contact on the body trunk, shoulder region and head of the demonstrator mouse), and body contact (measured as whole or partial body contact, initiated by the observer mouse, onset duration > 1 s). Non-social behaviors of observer mice during pre- and post-interaction tests, including tail rattling (characterized by the fast waving movements of the tail), self-grooming (defined as licking on the head or body trunk), immobility (defined as cessation of all movements and included sleeping, still and alert) and locomotion (other movement time except for specially defined actions).

Open-field testing used a transparent acrylic arena (40 × 40 × 40 cm). Mice explored freely for 5 minutes while total distance and center-versus-periphery occupancy were quantified with ANY-maze (Stoelting Co., v0.7.0.0). The arena was cleaned with 75% ethanol between trials.

Elevated-plus-maze testing used two open and two enclosed arms (30 × 7 cm; enclosed-wall height, 15 cm) elevated 40 cm above the floor. Mice explored for 5 minutes while behavior was recorded from above and analyzed with ANY-maze; open-arm time was the primary anxiety-related measure. The maze was cleaned with 75% ethanol between trials.

Behavioral state inference

Social-interaction videos were analyzed with identity-preserving multi-animal pose tracking (26,27) and behavioral segmentation tools including Keypoint-MoSeq (14). Videos were acquired from above at 1,920 × 1,080 pixels and 60 frames s-1 with fixed exposure, gain, white balance, focus and flicker-free illumination; spatial calibration converted pixels to centimeters. The final state-inference matrix contained one row per mouse and five columns: attack rate, tail-rattle duration, sniffing duration, immobility duration and self-grooming duration. No approach, retreat, interanimal-distance, spatial-occupancy or locomotor variable entered the final five-feature classifier.

For each mouse, the post-stress behavioral feature vector comprised attack rate, tail-rattle duration, sniffing duration, immobility duration and self-grooming duration, yielding one five-dimensional observation per mouse. Each feature was standardized across the discovery cohort as

$$z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j},$$

where x_{ij} is feature j for mouse i, and μ_j and σ_j are the discovery-cohort mean and standard deviation of that feature. Gaussian mixture models with diagonal covariance matrices and $K = 1$–5 components were fitted to the standardized mouse-level matrix using multiple random initializations, retaining the solution with the highest likelihood. Model complexity was assessed using the Bayesian information criterion,

$$\mathrm{BIC} = -2\ell(\hat{\theta}) + q\ln(n),$$

where $\ell(\hat{\theta})$ is the maximized log likelihood, q is the number of free model parameters and n is the number of mice. Generalization was assessed by fivefold cross-validation repeated ten times;

for each split, preprocessing and model fitting were performed using the training mice only, and performance was quantified as the mean held-out log likelihood,

$$\overline{\ell}_{\text{held-out}} = \frac{1}{n_{\text{test}}} \sum_{i=1}^{n_{\text{test}}} \log p\left(\mathbf{z}_i \mid \hat{\theta}_{\text{train}}\right).$$

Cluster stability was estimated from 100 mouse-level bootstrap resamples and quantified by the adjusted Rand index between the full-cohort solution and labels predicted by each bootstrap-refitted model after optimal label matching; the $K = 1$ adjusted Rand index was considered mathematically trivial and was not used to support model selection. The two-component solution was selected jointly on the basis of lower BIC, higher held-out likelihood and greater non-trivial bootstrap stability. Each mouse was assigned to the component with the largest posterior probability,

$$\hat{c}_i = \arg\max_k P(c_i = k \mid \mathbf{z}_i),$$

and biological state names were applied only after model fitting: the component with the higher mean attack rate was termed aggression-prone and the other avoidance-prone. Principal component analysis was applied to the same standardized matrix for visualization only and did not determine state assignment. The mouse-by-mouse correlation-distance matrix was defined as

$$d_{ij} = 1 - r_{ij},$$

where r_{ij} is the Pearson correlation between the five-feature vectors of mice i and j. Standardized feature profiles were reported as the state-specific mean z-score ± s.e.m., with

$$\text{s. e. m.} = \frac{s}{\sqrt{n}},$$

and exploratory between-state feature comparisons used two-sided tests at the mouse level.

For the baseline-to-post-stress trajectories in Figure 1H, the scaler and principal-component axes were fitted exclusively to the post-stress discovery data and then held fixed when transforming the matched baseline measurements. Each mouse's displacement in the standardized five-dimensional space was calculated as

$$\Delta_i = \left\| \mathbf{z}_{i,\text{post}} - \mathbf{z}_{i,\text{baseline}} \right\|_2.$$

Baseline separation between the two subsequently assigned states was quantified as the Euclidean distance between their baseline centroids and evaluated using 10,000 mouse-level label permutations. The permutation probability was calculated as

$$P = \frac{1 + \sum_{b=1}^{B} \mathbb{I}\,(D_b \geq D_{\mathrm{obs}})}{B + 1},$$

where D_{obs} is the observed centroid distance, D_b is the distance obtained after permutation b, and $B = 10{,}000$. For the independent replication cohort, the discovery-cohort feature means, standard deviations and fitted mixture parameters were frozen before application, and no model refitting or relabeling was performed.

For Figure 1J, baseline social, locomotor and anxiety indices were standardized and reduced to an unsupervised first principal-component score,

$$s_i = \mathbf{w}_1^{\top}\mathbf{z}_i,$$

without using later state labels during score construction. Discrimination of the subsequent state was summarized by the receiver-operating-characteristic area under the curve. Its 95% confidence interval was estimated from 5,000 stratified mouse-level bootstrap resamples, and deviation from chance was evaluated using 10,000 mouse-level label permutations. Throughout, the mouse was the biological replicate; trials, behavioral events and video frames were not treated as independent samples.

Stereotaxic surgery and activity mapping

Mice were anesthetized with isoflurane and placed in a stereotaxic frame. Viral vectors were delivered at 10 nL min-1 with a 15-min dwell time. Working coordinates relative to bregma were DP: AP +1.78 mm, ML ±0.30 mm, DV -3.50 mm; BNST: AP +0.38 mm, ML ±0.50 mm, DV -4.25 mm; and DMH: AP -1.70 mm, ML ±0.30 mm, DV -5.00 mm. Virus included AAV2/9-EF1a-Con Fon-GCaMP6s in DP; AAV2/2Retro-hSyn-Cre-WPRE-pA in BNST; AAV2/2Retro-hSyn-Flpo-WPRE-pA in DMH; and the Cre-, Flp- or dual-recombinase-dependent fluorescent, optogenetic and chemogenetic vectors specified for each experiment. Chronically implanted optical fibers had a 200-μm core and 0.37 NA.

Mice were anesthetized with isoflurane and placed in a stereotaxic frame. Viral vectors were delivered at 10 nL/min with a 15-minute dwell time. Working coordinates relative to bregma were DP: AP +1.78 mm, ML ±0.30 mm, DV -3.50 mm; BNST: AP +0.38 mm, ML ±0.50 mm, DV -4.25 mm; and DMH: AP -1.70 mm, ML ±0.30 mm, DV -5.00 mm. Placements were verified

histologically. Vectors included AAV2/9-EF1a-Con Fon-GCaMP6s (50 nL; 5e12 vg/mL; BC2016/9) in DP; AAV2/9-EF1α-DIO-EGFP-WPRE-pA (50 nL; 1.16e13 vg/mL; S1297-2), AAV2/1-hSyn-Cre-pA (30 nL; 5.78e12 vg/mL; S0292-1), AAV2/9-EF1α-fDIO-mCherry-WPRE-pA (50 nL; 1.65e13 vg/mL; S0553-9) as indicated; AAV2/2Retro-hSyn-Cre-WPRE-pA (30 nL; 2.05e13 vg/mL; S0278-2R), AAV2/9-GAD67-DIO-hM4Di (50 nL; >5e12 vg/mL; PT-0608), AAV2/2Retro-hSyn-Flpo-WPRE-pA (30 nL; 1.16e13 vg/mL; S0271-2R), AAV2/9-GAD67-DIO-hM3Dq (50 nL; >5e12 vg/mL; PT-3542) in BNST or DMH; and Cre-dependent ChR2 (AAV2/9-hEF1a-DIO-hChR2(H134R)-mCherry-WPRE-pA, 30 nL, >1e13 vg/mL; S0170-9), GtACR1 (AAV2/9-CAG-DIO-hGtACR1-P2A-EGFP-WPRE-pA, 30 nL, >1e13 vg/mL; S0311-9), and fluorescent-control vectors in BNST/DMH. ChR2 and GtACR1 were used for optical activation and inhibition, respectively (32,33). Chronically implantable optical fibers had a 200 μm core and 0.37 NA.

TRAP2;Ai9 mice were used for activity-dependent labeling (30). 4-hydroxytamoxifen (4-OHT; Aladdin, Cat. No. 10540-29-1) was prepared at 20 mg mL-1 in corn oil/ethanol (9:1) and administered at 75 mg kg-1 intraperitoneally immediately after the observation task. Matched-context controls underwent the same apparatus exposure without witnessed aggression. TRAP-positive cells were quantified from 3-5 sections per mouse, averaged within mouse and analyzed with mouse as the replicate.

Neural state decoding and output-bias recording

Dual systems dependent GCaMP virus AAV2/9-EF1a-Con Fon-GCaMP6s (29) was expressed in DP. Approximately 3 weeks later, a GRIN lens (ProView, 0.5-mm diameter, 4.89-mm length) was implanted above the targeted territory, and an nVista 3.0 or nVue 1.0 miniature microscope was baseplated after identification of the focal plane (28). Recorded cells were not retrogradely identified by projection target.

Miniscope movies were spatially downsampled, motion-corrected and segmented into neuronal regions of interest with codes from Peng Yuan's lab (45). Regions of interest were retained only when they met prespecified criteria for spatial morphology, signal-to-noise ratio, temporal stability and absence of motion contamination; overlapping or unstable components were removed before analysis.

Both sites from the same mouse were required for all bias analyses. The reference-corrected signal at each target $r \in \{BNST, DMH\}$ was first expressed as $\Delta F/F$:

$$x_{r,i}(t) = \frac{F_{r,i}(t) - \hat{F}_{0,r,i}(t)}{\hat{F}_{0,r,i}(t)} \quad (2)$$

To avoid direct comparison of raw fluorescence amplitudes across fibers, fluorophores or target sites, BNST and DMH signals were standardized independently within each mouse. The mean and standard deviation were calculated and then applied to the entire trace:

$$z_{r,i}(t) = \frac{x_{r,i}(t) - \mu_{r,i,R}}{\sigma_{r,i,R}} \quad (3)$$

The prespecified instantaneous output-bias signal for mouse i was defined as the difference between independently standardized target signals:

$$B_i(t) = z_{BNST,i}(t) - z_{DMH,i}(t) \quad (4)$$

Positive values indicate relatively greater BNST-directed activity, negative values indicate relatively greater DMH-directed activity and zero indicates equal standardized activity. This quantity is a relative, dimensionless signal and should not be interpreted as a comparison of absolute calcium amplitudes or axon numbers across sites.

For each event e in mouse i, the bias trace was aligned to event time ti,e. The mouse-level event-triggered average was computed before group averaging:

$$\bar{B}_i(\tau) = \frac{1}{N_i} \sum_{e=1}^{N_i} B_i(t_{i,e} + \tau) \quad (5)$$

Ni is the number of accepted events for mouse i, and τ is time relative to event onset. Events were included only when the complete analysis window was available.

The primary scalar mouse-level bias score was the mean event-aligned bias within the prespecified response window W = [τ1, τ2]:

$$M_i = \frac{1}{\tau_2 - \tau_1} \int_{\tau_1}^{\tau_2} \bar{B}_i(\tau) d\tau \quad (6)$$

For discretely sampled data, the integral was evaluated numerically using the sampling interval Δt. The exact response window -5-10 was fixed before examining state labels. Figure 4D displays the time-resolved mouse-level $\bar{B}i(\tau)$, whereas Figure 4E compares Mi between later aggression-prone and avoidance-prone mice. The prespecified scalar Mi was the primary endpoint to avoid treating time bins as independent observations.

As a complementary integrated measure, target-specific event-aligned activity was calculated for each mouse by numerical integration over the same response window:

$$\mathrm{AUC}_{r,i} = \int_{\tau_1}^{\tau_2} \bar{z}_{r,i}(\tau)\mathrm{d}\tau \quad (7)$$

$$\Delta\mathrm{AUC}_i = \mathrm{AUC}_{\mathrm{BNST},i} - \mathrm{AUC}_{\mathrm{DMH},i} \quad (8)$$

Here, Δt is the sampling interval after resampling, and $\bar{z}_{r,i}(\tau)$ is the event-averaged standardized signal at target r. A positive ΔAUC indicates greater integrated BNST-directed activity, whereas a negative value indicates greater integrated DMH-directed activity. Because division is unstable when the denominator approaches or crosses zero, BNST/DMH ratios should not be used as the primary complementary measure; ΔAUC is mathematically consistent with the difference-based bias definition in Eq. 4.

For state prediction in Figure 4F, the continuous mouse-level score Mi was entered directly into a receiver-operating-characteristic analysis, with the later behavioral state defined independently from post-stress behavior. The area under the curve was computed at the mouse level. No event, trial or time bin was counted as an independent sample.

For Figure 4G, one DP/DTT soma-state score and one projection-bias score were retained per mouse. Their association was quantified using Pearson correlation after visual inspection of linearity and influential observations; a Spearman rank correlation was used as a sensitivity analysis when appropriate. Correlation coefficients were calculated across mice, and the reported sample size was the number of animals with complete soma and dual-target recordings. The correlation does not establish that the somatic and terminal signals originated from the same dual-projecting neurons unless indicator expression was restricted to that population.

Fiber photometry

For dual-target photometry, fibers (Inper, Hangzhou, China) were positioned above BNST and DMH terminal fields in the same mouse. Signals therefore reflected all labeled DP axons present in each field and were not restricted to dual-projecting neurons. Data acquisition and normalization were performed with MATLAB. ΔF/F signals were z-scored, and calcium activity was quantified using the area under the curve (AUC). Statistical comparisons were made using parametric tests to evaluate event-related neural dynamics.

In vivo chemogenetic manipulation

Chemogenetic experiments were carried out following established protocols. Mice received bilateral injections of viral constructs expressing hM4Di, hM3Dq, or mCherry (control). Behavioral assays were performed 4 weeks post-surgery. Clozapine-N-oxide (CNO; LKT Labs,

Cat. No. 34233-69-7) was dissolved in sterile saline and freshly diluted before each use. For all experiments, mice received an intraperitoneal injection of CNO (5 mg/kg) 40 minutes before the test stage.

Imaging and quantification

Mice were perfused, fixed-brains were cryoprotected in 30% sucrose. Frozen brains were sliced into 40-μM sections using a cryostat (Leica, CM1950, Germany). Sections were then stained with DAPI, and washed with PBS, transferred onto Super Frost slides, and mounted under glass coverslips with mounting media.

Imaging was performed using VS200 (Olympus, Japan) or FV3000 confocal microscopes (Olympus, Japan) and was processed and analyzed using ImageJ software v.1.51j8 (NIH). For Fos-TRAP+ cell quantifications, images were performed using the 'analyze particles' function in each channel, followed by counting the intersection of the particles for each region of interest. Averages were collected from 3–5 brain slices per mouse. For viral expression, images were background subtracted and then thresholded using the Max Entropy function. For photometry validation, we confirmed before behavior testing that GCaMP6 signal levels in all animals varied independently from the signal change in the auto fluorescence channel. After behavior testing, we confirmed that optic fiber placement was accurate. If their viral expression could not be validated or when there was equipment failure during behavior, mice were excluded from the analysis.

Statistical analysis

All tests were two sided unless a directional test was prespecified. Mouse was the biological replicate; neurons, trials and sections were nested within mouse. Repeated measures were analyzed with linear or generalized linear mixed-effects models when appropriate. Behavioral-state model selection used BIC, held-out log likelihood and bootstrap adjusted Rand index. Decoder performance used nested leave-one-mouse-out cross-validation and mouse-level label permutation. State probability denotes GMM posterior probability; categorical state assignment denotes the maximum-posterior class; cohort state proportion denotes the fraction assigned to each class. Correlations used one mouse-level summary per animal. Multiple behavioral-control outcomes were corrected using the prespecified multiplicity procedure.**Results**

Result

Witnessed social stress drives divergent social states

To investigate how a common stressor is followed by divergent social strategies, we used a witnessed social stress assay that separated observation of repeated aggression from direct physical defeat. Observers remained behind a transparent perforated barrier while a resident aggressor attacked a conspecific victim (**Fig. 1A**), and each observer underwent standardized free-interaction tests before and after the observation period (**Fig. 1B**). Witnessed attacks elicited acute heart-rate responses (**Fig. 1C**), confirming physiological engagement. Aggregate heart-rate measures did not differ between mice subsequently assigned to the aggression-prone and avoidance-prone states (**Fig. S2C**; n = 22 and 26, respectively), arguing against a gross difference in acute autonomic engagement as the sole explanation for later divergence.

Post-stress social behavior varied markedly across animals. Aggression-prone mice showed higher attack rate, tail-rattle duration and social sniffing, whereas separation in immobility and self-grooming was weaker (**Fig. S2D-F**; attack and tail rattling, $p < 0.0001$; sniffing, $p = 0.0001$; immobility, $p = 0.0758$; self-grooming, $p = 0.3451$; n = 22 and 26). Avoidance-prone mice also showed lower open-field center occupancy ($p = 0.001$), whereas elevated-plus-maze behavior showed weaker separation ($p = 0.5337$). Sucrose preference and social preference did not differ detectably between states (**Fig. S2B,D-E**; $p = 0.7381$ and $p = 0.5749$, respectively). Thus, the most pronounced post-stress variation occurred in the social-behavior domain.

Given that post-stress behavioral variation was predominantly social and not easily captured by single metrics, we adopted a data-driven approach rather than relying on a predefined behavioral threshold to classify susceptibility. We constructed comprehensive mouse-level feature vectors encompassing social events, pose dynamics, spatial variables, and locomotion (**Fig. 1D**). After z-standardization, diagonal-covariance Gaussian mixture models containing one to five components were compared. A two-component model had the lowest Bayesian information criterion (BIC = 691.75), the best held-out mean log likelihood (-7.05) and the highest nontrivial bootstrap stability (**Fig. 1E,F and Fig. S3**; adjusted Rand index = 0.64). Principal-component analysis and the mouse-by-mouse correlation-distance matrix were used only as complementary visualizations. Post hoc feature profiles identified 22 aggression-prone and 26 avoidance-prone mice; state names were assigned only after unsupervised fitting (**Fig. 1G**). These distinct behavioral profiles led us to term them aggression-prone and avoidance-prone states, respectively, a nomenclature applied only after the unbiased model fitting.

The two post-stress states were not readily explained by baseline behavior. When matched baseline measurements were projected into the frozen post-stress feature space, the future groups substantially overlapped before observation (baseline centroid permutation $p = 0.88$), whereas within-mouse trajectories subsequently diverged toward distinct post-stress regions (**Fig. 1H**). A frozen five-feature pipeline was also applied to an independent cohort (**Fig. 1I**). Baseline social, locomotor and anxiety indices overlapped between the later states (**Fig. S1A-D**; $n = 22$ and 26), and a label-blind baseline principal-component score provided weak discrimination (**Fig. 1J**; AUC = 0.55; 95% bootstrap CI, 0.39-0.72; permutation $p = 0.52$). These findings collectively demonstrate that post-stress divergence is more consistent with an experience-dependent state transition, shaped by the stressor itself, rather than a simple expression of a pre-existing, immutable behavioral trait.

DP projections to BNST and DMH overlap anatomically

To examine how a common prefrontal population may support divergent social strategies, we first mapped the projection architecture of the DP. Anterograde tracing revealed dense DP axonal innervation in both the bed nucleus of the stria terminalis (BNST) and the dorsomedial hypothalamus (DMH) (**Fig. 2A**). Dual-retrograde tracing from the two targets showed that 61.28 ± 4.20% of labeled DP projection neurons were double labeled ($n = 10$ mice), whereas smaller fractions were BNST only (19.84 ± 2.80%) or DMH only (**Fig. 2B,C**; 18.87 ± 2.50%; one-way ANOVA, $p < 0.001$). Thus, DP neurons projecting to the BNST and DMH predominantly comprise a shared output population, rather than two anatomically segregated projection-defined cell classes (**Fig. 2D**).

Together, these findings indicate that BNST- and DMH-directed pathways are predominantly generated by an overlapping DP neuronal population rather than by distinct projection-defined cell classes. This shared-output architecture suggests that divergent social strategies could arise through target-specific allocation of activity within a common prefrontal population. We therefore next examined whether DP population activity diverges during stress observation, before the emergence of post-stress behavioral phenotypes.

DP activity during stress observation precedes the later social state

We next asked whether activity in the broader DP population diverged during observation, before post-stress behavior was expressed. Activity-dependent TRAP2 labeling showed increased DP recruitment after witnessed aggression (**Fig. 3A**). Bulk GCaMP6s recordings showed event-

related responses around witnessed attacks (**Fig. 3B**). To resolve these dynamics at cellular resolution, we performed miniscope calcium imaging of DP neurons during stress observation (**Fig. 3C,D**). Having established that DP projections to the BNST and DMH arise predominantly from a shared neuronal population, we injected retro-Cre in BNST and retro-Flpo in DMH, next examined whether DP activity diverges during stress observation, before the expression of post-stress social phenotypes.

DP population activity differed between mice that later entered aggression-prone and avoidance-prone states. Differences were most evident around attack onset and victim escape and were smaller during matched sniffing epochs (**Fig. 3E**). Low-dimensional representations of DP population activity further revealed separable neural trajectories for the two future behavioral states (Fig. 3F; 120 aggression-prone-associated and 120 avoidance-prone-associated neurons). The distance between these trajectories increased progressively over the observation session (Fig. 3G; n = 16). Thus, DP population dynamics diverged during stress observation and preceded the expression of post-stress social phenotypes.

A cross-animal decoder trained on observation-period DP activity became more informative over time (**Fig. 3H**; n = 30 mice) and predicted the subsequent state with AUC = 0.73 (**Fig. 3I**). Neural activity outperformed models based on baseline behavior, locomotion and combined non-neural variables (**Fig. 3J**; n = 24). Predictive information remained after regression of measured motor and autonomic covariates (**Fig. 3K**; n = 24), and performance exceeded a mouse-level label-permutation distribution (**Fig. 3L**; 500 permutations, p = 0.002). These findings indicate that DP population activity during stress observation contains information predictive of the subsequent social phenotype that is not explained by the measured movement and autonomic variables.

Together, these findings show that activity within the shared DP population diverges during stress observation and predicts the subsequent social state. We next investigated how this common prefrontal population generates distinct downstream consequences by simultaneously measuring DP output to the BNST and DMH in individual mice.

Relative BNST- and DMH-directed DP terminal activity predicts later social strategy

Having found that the shared DP population diverges during observation, we next asked whether this divergence is expressed as a relative bias in output to BNST versus DMH. We recorded DP axonal calcium signals simultaneously in BNST and DMH using dual-site fiber photometry (**Fig. 4A,B**). Event-aligned terminal activity differed according to the later behavioral

state (**Fig. 4C**). Total terminal activity did not differ between future states, whereas the relative bias did, suggesting that the shared population distributes its output differentially across targets rather than simply increasing or decreasing overall activity.

To quantify relative target engagement without directly comparing raw fluorescence amplitudes, each channel was standardized independently and the output-bias signal was defined as B = Z_BNST - Z_DMH. A complementary analysis used the difference in integrated event-aligned activity (**Fig. S4**; BNST AUC - DMH AUC). Future aggression-prone mice developed positive BNST-dominant bias, whereas future avoidance-prone mice developed negative DMH-dominant bias (**Fig. 4D**; n = 14 mice per state). Mouse-level mean bias differed between the displayed aggression-prone and avoidance-prone subsets (**Fig. 4E**; n = 7 and 6; two-sided t-test, $t = 7.77$, $p < 0.0001$).

Output bias predicted the later state with AUC = 0.71 (**Fig. 4F**). Projection-bias scores also correlated with DP soma-state scores across animals (**Fig. 4G**; n = 26, Pearson $r = 0.52$, $p = 0.0062$). These observations link a DP population signal to the relative activity of two terminal fields and are compatible with an output-bias model. They do not distinguish whether the bias arises within dual-projecting neurons, through differential recruitment of single-target neurons or through both mechanisms.

These findings raised the question of causality. If output bias contributes to the later state, manipulating the relative influence of the two downstream pathways should shift state probability. We therefore next manipulated DP-defined target populations in BNST or DMH during witnessed stress.

Manipulation of DP-input-defined downstream populations biases social-state outcomes

To test whether the association between DP output bias and subsequent behavior is causal, AAV1-hSyn-Cre was delivered to DP and Cre-dependent ChR2 or GtACR1 was expressed in BNST or DMH, thereby labeling populations associated with DP input (**Fig. 5A-C**). Activation of the labeled BNST population increased aggression-prone state probability, whereas inhibition reduced it (**Fig. 5D,E**; n = 14 in each displayed comparison; $p = 0.014$ and $p = 0.024$). Manipulation of the labeled DMH population produced complementary changes in avoidance-prone state probability (**Fig. 5F,G**; n = 14; $p = 0.0101$ and $p = 0.0424$). Across these perturbations, transition-matrix analysis revealed coordinated redistribution among neutral, aggression-prone and avoidance-prone regions of the behavioral state space (**Fig. 5H**).

We next assessed whether these pathways influence social-state formation during stress observation or the subsequent expression of an established state. Manipulation during witnessed stress produced a larger absolute classifier-score shift than manipulation restricted to the later social interaction (**Fig. 5I**; n = 28; p = 0.0094). This timing dependence is consistent with a contribution during state formation but does not exclude an effect on pre-existing bias or later expression. Locomotion, speed, object exploration and nonsocial anxiety measures showed no consistent pathway-general change (**Fig. 5J and Fig. S5**). Nominal effects in individual control measures did not survive the prespecified multiple-comparison procedure.

Together, these experiments support a contribution of the labeled BNST and DMH populations to behavioral-state outcomes.

Cell-type-specific manipulations shift established output bias and behavior

To identify the downstream cell types that might implement target-biased output, we focused on literature-guided candidates. BNST circuits are predominantly GABAergic and stress sensitive (12,13,34,35), whereas DMH receives DP stress-related input and contains functionally important vGlut2-positive glutamatergic populations (10,15). Labeling associated with the DP injection was enriched in GABAergic BNST neurons and vGlut2-positive DMH neurons (**Fig. 6A,B**), identifying candidate populations for manipulation while not establishing that all labeled cells were monosynaptic postsynaptic targets.

In aggression-prone mice, inhibition of GABAergic BNST neurons or activation of vGlut2-positive DMH neurons shifted the BNST-DMH bias index toward less BNST-dominant values and reduced the aggression-prone behavioral-state score (**Fig. 6C-E**; n = 15 and 16). The corresponding locomotor and nonsocial-anxiety controls showed no consistent change (**Fig. S6**). Because the manipulations targeted cell classes rather than DP synapses onto those cells, they establish experimental modifiability of the downstream state variable but not a DP-specific synaptic mechanism.

Collectively, the data support a working model in which overlapping DP projection anatomy coexists with state-predictive DP activity and opposing relative signals in BNST- and DMH-directed terminal fields (**Fig. 6F**). A BNST-dominant population-level bias is associated with aggression-prone outcomes, whereas a DMH-dominant bias is associated with avoidance-prone outcomes. Direct demonstration that the same dual-projecting neurons differentially regulate their two collateral fields will require intersectional recording restricted to that population.

Discussion

This study links witnessed social stress to reproducible post-stress behavioral states, prospective DP population activity, relative BNST- versus DMH-directed terminal signals and behavioral effects of downstream manipulations. The evidence supports target-biased population activity as a predictor of later strategy and shows that labeled downstream populations can influence classifier-derived state measures. It does not directly establish that a single dual-projecting DP population generates the two terminal signals or that differential regulation occurs between collateral branches of the same neurons.

The behavioral results extend work showing that observed aggression and direct social defeat can produce persistent but heterogeneous changes in later behavior (4,21-25,36-41). The two-state solution was supported by BIC, held-out likelihood and bootstrap stability, and baseline behavior showed weak discrimination of later assignment. Nevertheless, weak baseline prediction is not equivalent to absence of pre-existing differences. Intrinsic excitability, unmeasured behavioral traits and prior testing could all shape the later outcome.

The causal interpretation of the downstream manipulations is also limited by the targeting strategy. AAV1-associated labeling identifies candidate input-defined recipient populations but does not isolate DP synaptic transmission, because recipient neurons integrate other afferents and AAV1 can exhibit retrograde transport (44). Accordingly, activation or inhibition of the labeled BNST and DMH populations demonstrates their contribution to behavior but should not be described as direct manipulation of BNST-projecting or DMH-projecting DP neurons. The shared-population output-bias scheme is therefore presented as a working circuit model rather than an established cellular routing mechanism.

The projection architecture provides a circuit-level basis for this divergence. BNST- and DMH-projecting DP neurons overlapped extensively, indicating that the two outputs are predominantly generated by a shared prefrontal population (**Fig. 4B–D**). During stress observation, however, terminal activity in these pathways developed a target-specific bias that predicted the subsequent behavioral state. This organization supports a model in which a common DP population differentially weights its output across downstream targets. In this framework, the relative influence of BNST and DMH channels, rather than recruitment of fully segregated projection-defined cell classes, determines the direction of the post-stress social response. The anatomical

position of DP at the interface of cortical, limbic, hypothalamic and extended-amygdala networks is compatible with such a role in coordinating stress-related behavioral allocation (8-11).

Causal manipulations further indicate that this balance is established during stress exposure. Increasing or decreasing activity in DP pathways to the BNST and DMH shifted social-state probabilities in opposing directions. Moreover, perturbation during witnessed stress produced larger state shifts than the same manipulation delivered during the later social interaction (**Fig. 5I**). These timing-dependent effects indicate that DP downstream pathways contribute to state formation while the stressful event is being observed, rather than exclusively regulating the expression of an established behavioral state. More broadly, these results suggest that stress-related social behavior can emerge from changes in the relative engagement of interconnected circuit outputs rather than from transmission of fixed action commands through a single pathway (16-20).

Cell-type-specific experiments showed that an established post-stress state remains modifiable. In aggression-prone mice, inhibition of the GABAergic BNST branch or activation of the Vglut2-positive DMH branch shifted the BNST–DMH output-bias index towards balance and reduced the aggression-prone behavioral-state score (**Fig. 6C–E**). Thus, the circuit variable associated with state formation can be modified after the state has emerged. This result demonstrates circuit-level reversibility within the present model. It does not, however, establish a therapeutic target or demonstrate treatment of a psychiatric disorder (2,3,34).

Several limitations should be considered. First, witnessed aggression is a multimodal social experience, and the present experiments did not isolate the relative contributions of visual, auditory and olfactory cues. Second, the observed neural and behavioral responses do not establish empathy or attribution of the aggressor's or victim's internal state. Third, aggression and avoidance may each represent context-dependent responses to threat; their relevance to maladaptive behavior would require evidence of persistence, inappropriate generalization or functional cost. Finally, conclusions regarding sex-dependent mechanisms should be restricted to cohorts designed and powered to test sex as a biological variable.

In summary, these findings support a model in which individual divergence after a common social stressor arises from target-biased output dynamics within a shared DP population. Relative DP output to the BNST and DMH sets the balance between aggression-prone and avoidance-prone states, emerges during stress observation and remains modifiable through cell-type-specific rebalancing of downstream pathways (**Fig. 6F**). By integrating data-driven behavioral phenotyping,

prospective neural prediction, projection-specific output measurements, causal perturbation and circuit rebalancing, this study provides a mechanistic framework for stress-related individual differences without assuming that vulnerability is fixed before exposure.

Figure legends

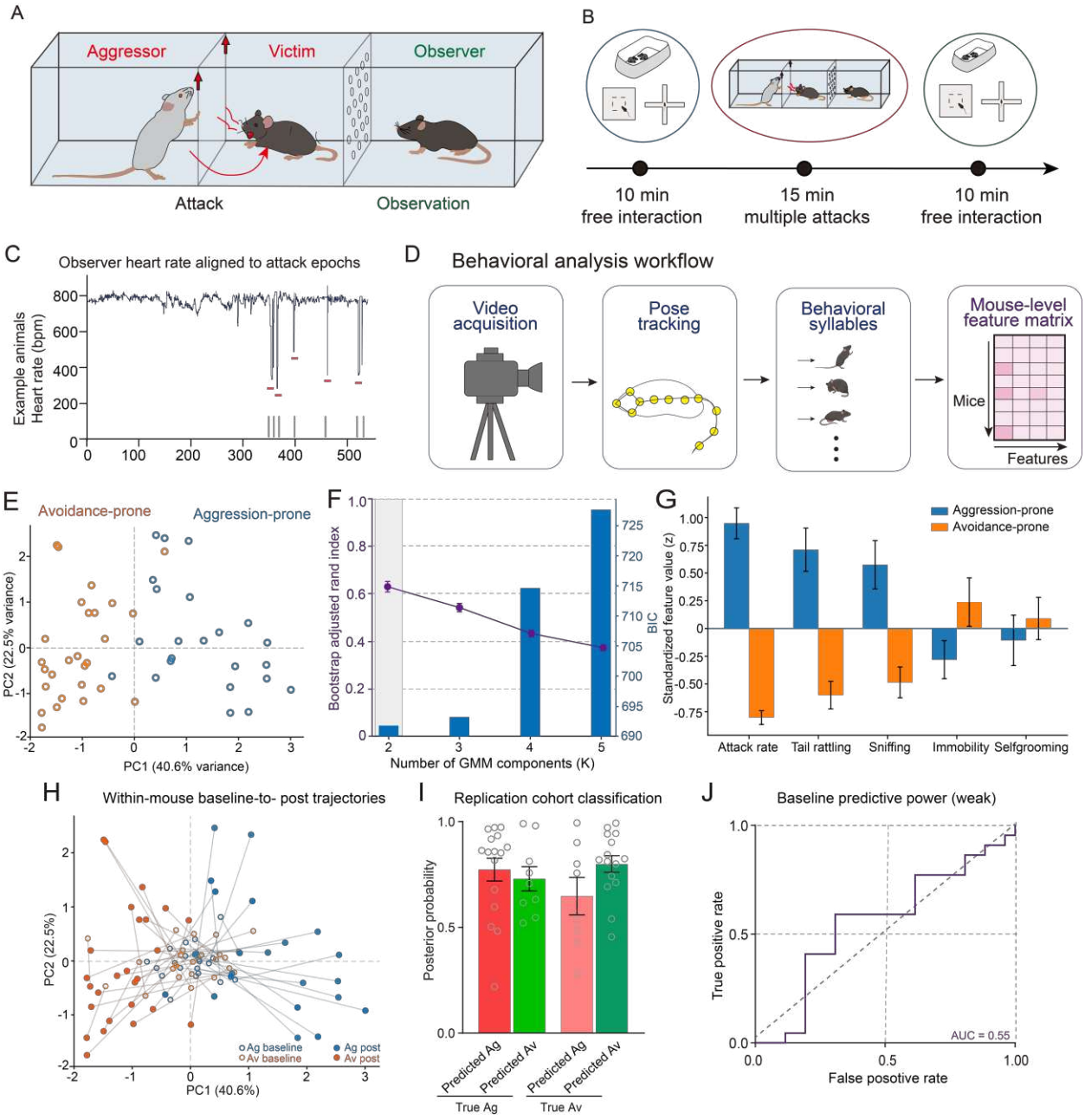


Fig. 1 Witnessed social stress reveals divergent post-stress social states. (A) Witnessed-aggression assay. (B) Baseline interaction, 15-min observation and post-stress interaction. (C) Representative observer heart-rate trace aligned to witnessed attacks. (D) Five-feature behavioral-analysis workflow. (E) PCA visualization of the post-stress feature matrix, colored only after

unsupervised state assignment (n = 22 aggression-prone; n = 26 avoidance-prone). (F) Comparison of one- to five-component GMMs by BIC, held-out likelihood and bootstrap stability. (G) Standardized post-stress profiles: attack rate, $p < 0.0001$; tail rattling, $p < 0.0001$; sniffing, $p = 0.00011$; immobility, $p = 0.0758$; self-grooming, $p = 0.506$. (H) Matched baseline measurements projected into the frozen post-stress feature space, with within-mouse baseline-to-post trajectories (n = 22 and 26; baseline centroid permutation $p = 0.88$). (I) Independent-cohort posterior probabilities under the frozen classifier. (J) ROC curve for the label-blind baseline PC1 score (AUC = 0.55; 95% bootstrap CI, 0.39-0.72; permutation $p = 0.52$).

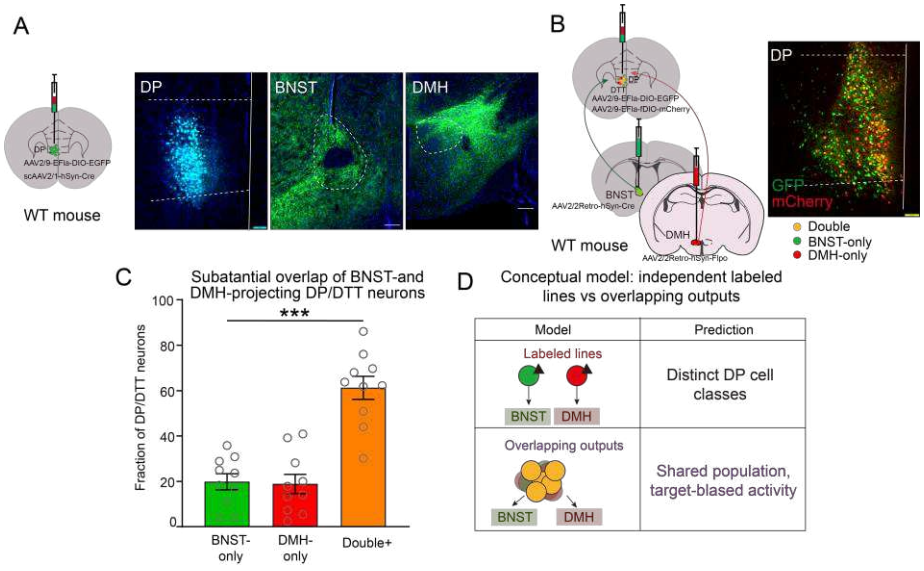


Fig. 2 DP projections to BNST and DMH show substantial anatomical overlap. (A) Anterograde DP labeling and representative axonal fields in BNST and DMH. (B) Dual-retrograde labeling from BNST and DMH. (C) Fractions of BNST-only, DMH-only and double-labeled DP neurons (n = 10 mice; 19.84 ± 2.80%, 18.87 ± 2.50% and 61.28 ± 4.20%, respectively). (D) Anatomical working model.

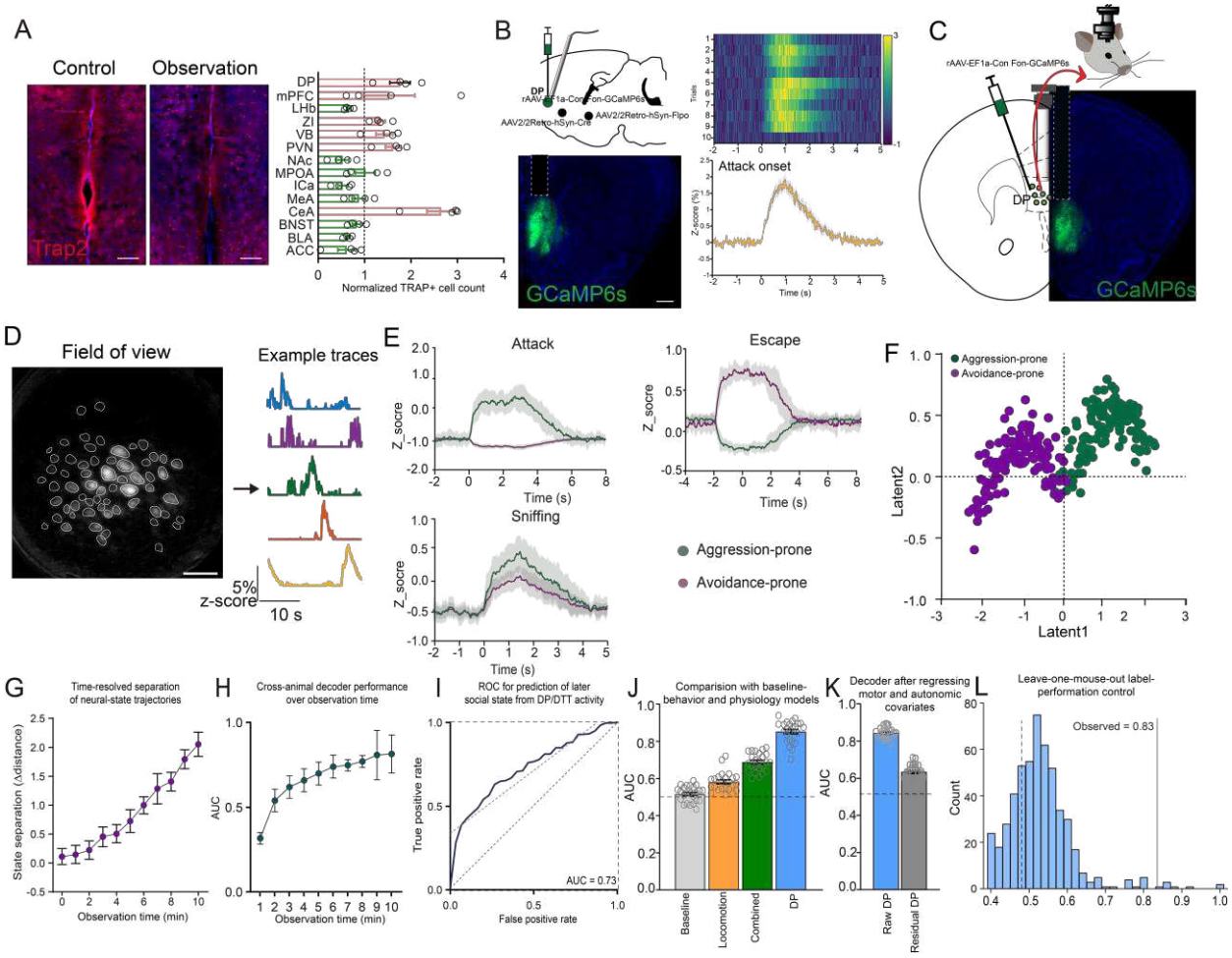


Fig. 3 DP population dynamics during stress predict the later social state. (A) Representative TRAP2 labeling and normalized regional counts after matched context or witnessed stress (red indicates the increased brain areas, green indicates the decreased brain areas). (B) DP bulk-GCaMP6s recording strategy, event heat map and mean response aligned to witnessed attacks. (C) Viral strategy and representative GCaMP6s expression for miniscope imaging. (D) Field of view, segmented cells and example traces. (E) Event-aligned activity around attack onset, victim escape and matched sniffing. Time zero denotes event onset. (F) Descriptive cell-level embedding (120 displayed cells per future-state group); cells were not projection identified and mouse was the inferential unit. (G) Time-resolved separation of mouse-level future-state trajectories (n = 16). (H) Cross-animal decoder performance over observation (n = 30). (I) ROC prediction of later state (AUC = 0.73). (J) Comparison with baseline, locomotor and combined non-neural models (n = 24). (K) Performance before and after regression of motor/autonomic covariates (n = 24). (L) Mouse-level label-permutation distribution (500 permutations; p = 0.002; n = 24).

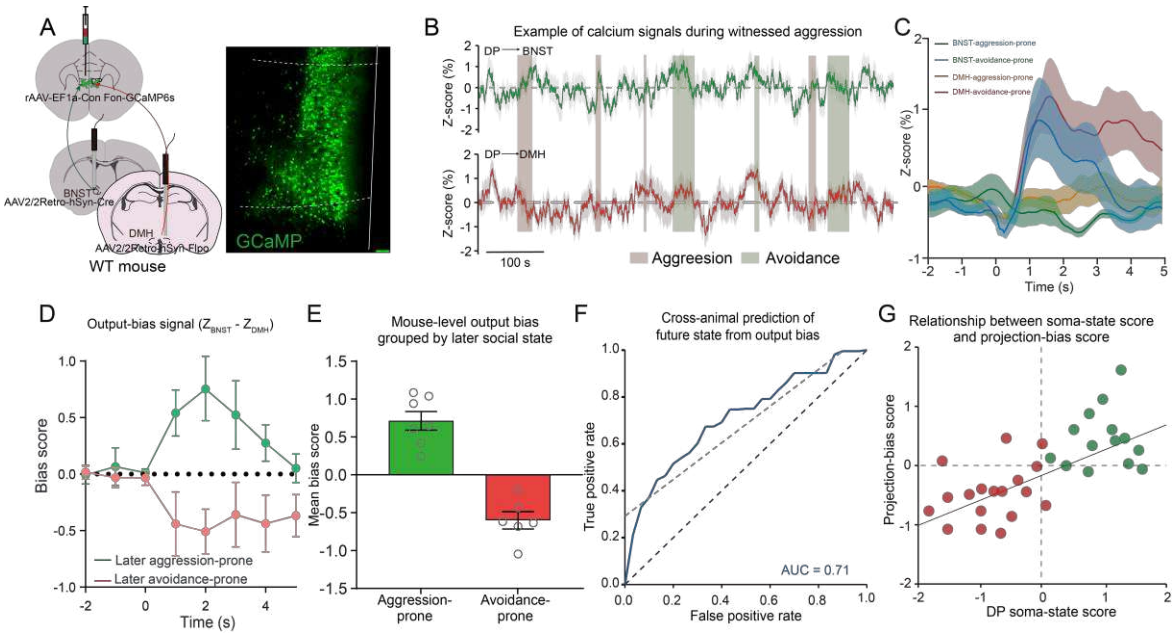


Fig. 4 Relative BNST- and DMH-directed DP terminal activity predicts later social strategy. (A) Simultaneous recording from DP axonal fields in BNST and DMH. The recordings were not restricted to dual-projecting neurons. (B) Reference-corrected terminal signals. (C) Event-aligned BNST- and DMH-directed activity grouped by later state. (D) Output-bias signal B = Z_BNST - Z_DMH (n = 14 mice per state). (E) Mouse-level mean bias (n = 7 aggression-prone; n = 6 avoidance-prone; $t(11) = 7.77$, $p < 0.0001$). (F) ROC prediction from output bias (AUC = 0.71). (G) Correlation between DP soma-state and projection-bias scores (n = 26; Pearson r = 0.52, $p = 0.0062$).

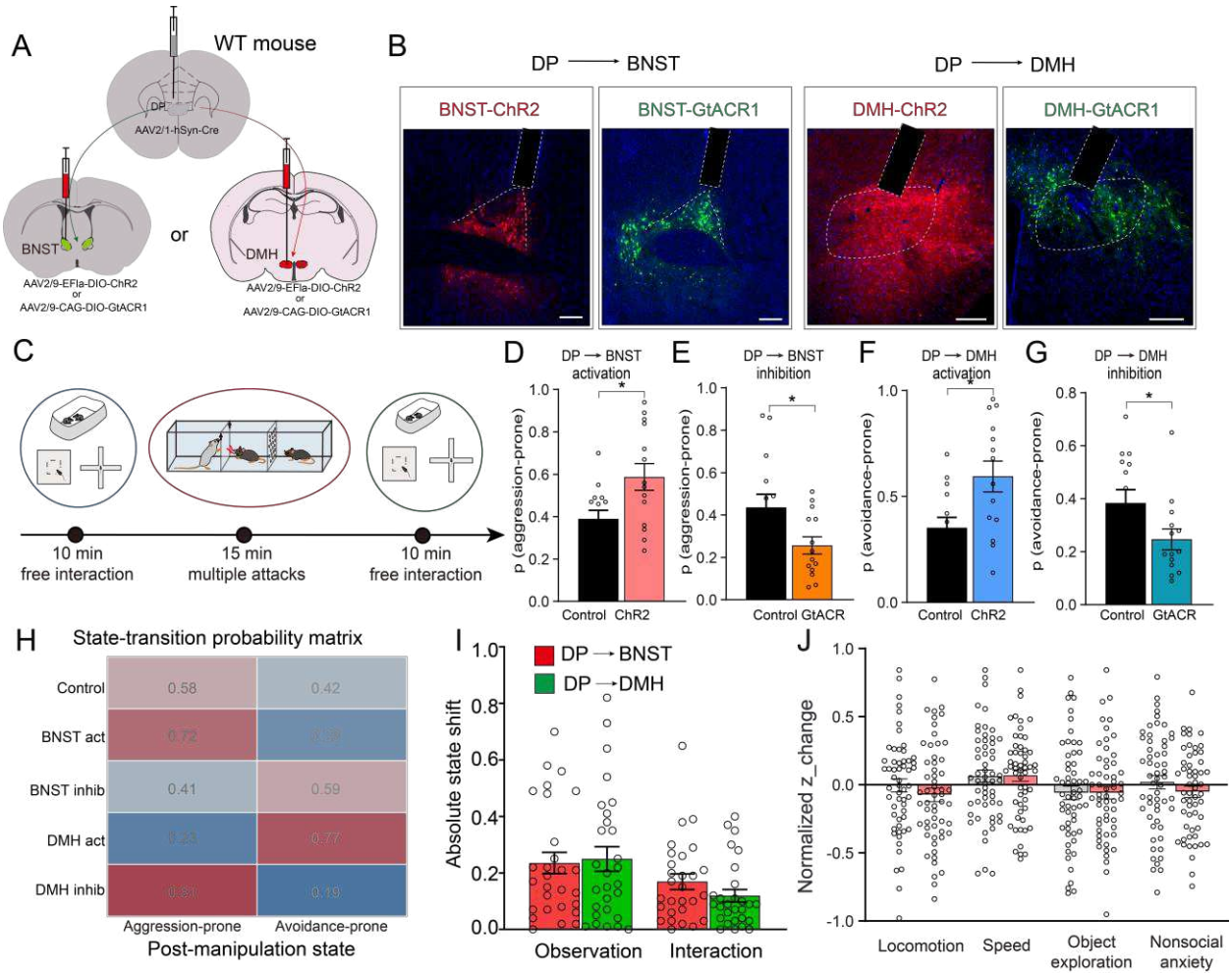


Fig. 5 Manipulation of DP-input-defined downstream populations biases social-state outcomes. (A) Virus injection strategy for labeling BNST or DMH populations associated with DP input and expression of ChR2 or GtACR1. (B) Representative expression and fiber placements. (C) Behavioral timeline. (D,E) Activation or inhibition of the labeled BNST population and aggression-prone posterior probability (n = 14; p = 0.014 and p = 0.024). (F,G) Activation or inhibition of the labeled DMH population and avoidance-prone posterior probability (n = 14; p = 0.0101 and p = 0.0424). (H) Distribution of classifier assignments after manipulation. (I) Absolute classifier-score shift for manipulation during observation versus the later interaction (n = 28; p = 0.0094). (J) Normalized z_socre of locomotion, speed, object exploration and nonsocial-anxiety controls (n = 56).

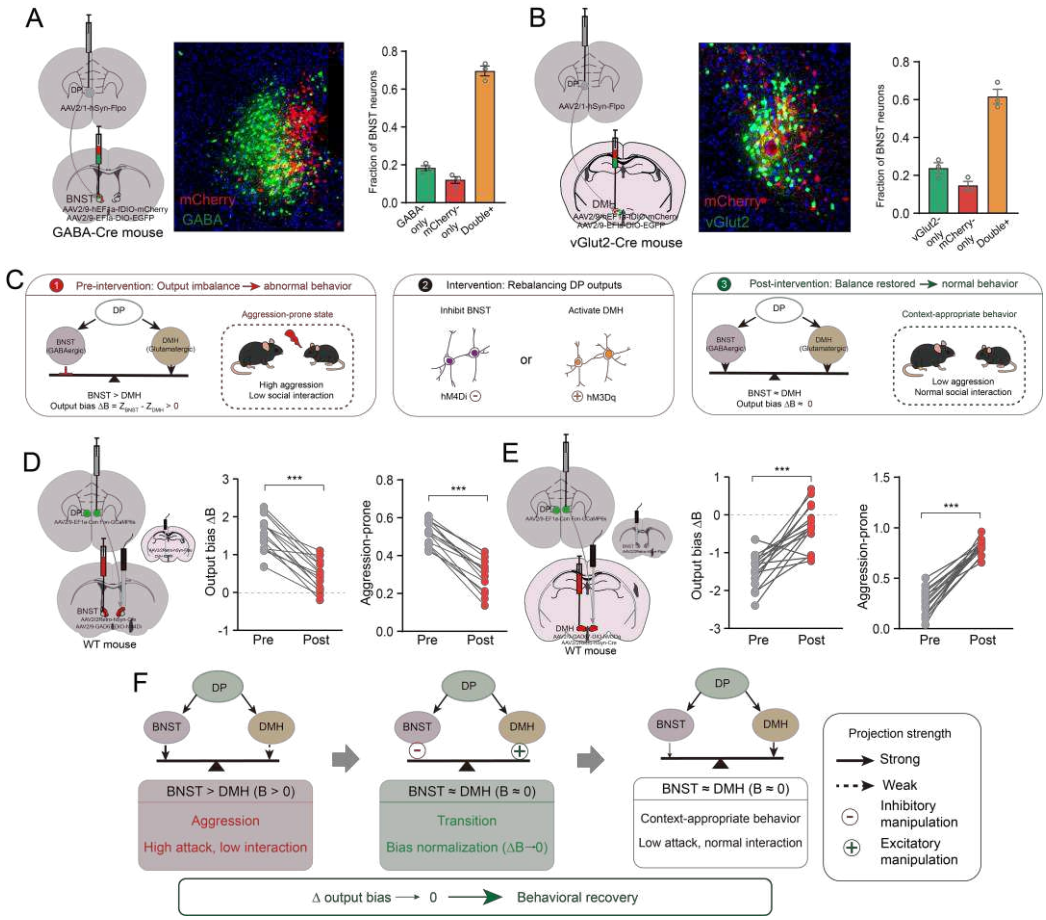


Fig. 6 Cell-type-specific downstream manipulations shift output bias and behavior. (A) Strategy of virus injection in GABAergic BNST neurons associated with the DP injection. (B) Strategy of virus injection in vGlut2-positive DMH neurons. (C) Experimental strategy for inhibiting BNST GABAergic neurons or activating DMH vGlut2-positive neurons after the aggression-prone state was established. (D) Paired output-bias and behavioral-state scores after BNST GABAergic inhibition (n = 15). (E) Paired output-bias and behavioral-state scores after DMH vGlut2-positive activation (n = 16). (F) Working model linking overlapping DP projection anatomy, relative terminal-field activity and downstream behavioral effects. The model is not direct evidence for collateral-specific regulation by dual-projecting neurons.

Acknowledgments

This work was supported by Fudan University (JIH2641050Y to H.H.).

Data and Code Availability

All data and code needed to evaluate and reproduce the results in the paper are present in the paper and/or the Supplementary Materials.

Disclosures

The authors report no biomedical financial interests or potential conflicts of interest.

Author Contributions

H.H., W.W. and J.Z. conceived the study; H.H. and W.R. developed the methodology; H.H., W.R., L.L. and Z.L. performed behavioral tests, viral injection, and imaging; W.R. performed the computational analysis; H.H. wrote the main manuscript text; W.W. revised the manuscript; H.H. acquired funding. All authors reviewed the manuscript.

References

1. Willmore, L., Cameron, C., Yang, J., Witten, I. B. & Falkner, A. L. Behavioural and dopaminergic signatures of resilience. Nature 611, 124-132 (2022). https://doi.org/10.1038/s41586-022-05328-2.

2. Nestler, E. J. & Russo, S. J. Neurobiological basis of stress resilience. Neuron 112, 1911-1929 (2024). https://doi.org/10.1016/j.neuron.2024.05.001.

3. Sanacora, G., Yan, Z. & Popoli, M. The stressed synapse 2.0: pathophysiological mechanisms in stress-related neuropsychiatric disorders. Nat. Rev. Neurosci. 23, 86-103 (2022). https://doi.org/10.1038/s41583-021-00540-x.

4. Schuler, H., Eid, R. S., Wu, S., Tse, Y. C., Cvetkovska, V., Lopez, J., Quinn, R., Zhou, D., Meccia, J., Dion-Albert, L., Bennett, S. N., Newman, E. L., Trainor, B. C., Peña, C. J., Menard, C., & Bagot, R. C. Data-driven analysis identifies novel modulation of social behavior in female mice witnessing chronic social defeat stress. Biol. Psychiatry 98, 416-426 (2025). https://doi.org/10.1016/j.biopsych.2024.11.017.

5. Battivelli, D., Fan, Z., Hu, H. & Gross, C. T. How can ethology inform the neuroscience of fear, aggression and dominance? Nat. Rev. Neurosci. 25, 809-819 (2024). https://doi.org/10.1038/s41583-024-00858-2.

6. Tseng, Y. T., Schaefke, B., Wei, P., & Wang, L. Defensive responses: behaviour, the brain and the body. Nat. Rev. Neurosci. 24, 655-671 (2023). https://doi.org/10.1038/s41583-023-00736-3.

7. LeDoux, J. & Daw, N. D. Surviving threats: neural circuit and computational implications of a new taxonomy of defensive behaviour. Nat. Rev. Neurosci. 19, 269-282 (2018). https://doi.org/10.1038/nrn.2018.22.

8. Botterill, J. J., Khlaifia, A., Appings, R., Wilkin, J., Violi, F., Premachandran, H., Cruz-Sanchez, A., Canella, A. E., Patel, A., Zaidi, S. D., & Arruda-Carvalho, M. Dorsal peduncular cortex activity modulates affective behavior and fear extinction in mice. Neuropsychopharmacology 49, 993-1006 (2024). https://doi.org/10.1038/s41386-024-01795-5.

9. Nakamura, K., Nakamura, Y. & Kataoka, N. A hypothalamomedullary network for physiological responses to environmental stresses. Nat. Rev. Neurosci. 23, 35-52 (2022). https://doi.org/10.1038/s41583-021-00532-x.

10. Kataoka, N., Shima, Y., Nakajima, K. & Nakamura, K. A central master driver of psychosocial stress responses in the rat. Science 367, 1105-1112 (2020). https://doi.org/10.1126/science.aaz4639.

11. Wang, H., Wang, Q., Cui, L., Feng, X., Dong, P., Tan, L., Lin, L., Lian, H., Cao, S., Huang, H., Cao, P., & Li, X. M. A molecularly defined amygdala-independent tetra-synaptic forebrain-to-hindbrain pathway for odor-driven innate fear and anxiety. Nat. Neurosci. 27, 514-526 (2024). https://doi.org/10.1038/s41593-023-01562-7.

12. Hammack, S. E., Braas, K. M. & May, V. Chemoarchitecture of the bed nucleus of the stria terminalis: neurophenotypic diversity and function. Handb. Clin. Neurol. 179, 385-402 (2021). https://doi.org/10.1016/B978-0-12-819975-6.00025-X.

13. Partridge, J. G., Forcelli, P. A., Luo, R., Cashdan, J. M., Schulkin, J., Valentino, R. J., & Vicini, S. Stress increases GABAergic neurotransmission in CRF neurons of the central amygdala and bed nucleus stria terminalis. Neuropharmacology 107, 239-250 (2016). https://doi.org/10.1016/j.neuropharm.2016.03.029.

14. Weinreb, C., Pearl, J. E., Lin, S., Osman, M. A. M., Zhang, L., Annapragada, S., Conlin, E., Hoffmann, R., Makowska, S., Gillis, W. F., Jay, M., Ye, S., Mathis, A., Mathis, M. W., Pereira, T., Linderman, S. W., & Datta, S. R. Keypoint-MoSeq: parsing behavior by linking point

tracking to pose dynamics. Nat. Methods 21, 1329-1339 (2024). https://doi.org/10.1038/s41592-024-02318-2.

15. Francois, M., Kaiser, L., He, Y., Xu, Y., Salbaum, J. M., Yu, S., Morrison, C. D., Berthoud, H. R., & Münzberg, H. Leptin receptor neurons in the dorsomedial hypothalamus require distinct neuronal subsets for thermogenesis and weight loss. Metabolism 163, 156100 (2025). https://doi.org/10.1016/j.metabol.2024.156100.
16. Yang, B., Karigo, T. & Anderson, D. J. Transformations of neural representations in a social behaviour network. Nature 608, 741-749 (2022). https://doi.org/10.1038/s41586-022-05057-6.
17. Minakuchi, T., Guthman, E. M., Acharya, P., Hinson, J., Fleming, W., Witten, I. B., Oline, S. N., & Falkner, A. L. Independent inhibitory control mechanisms for aggressive motivation and action. Nat. Neurosci. 27, 702-715 (2024). https://doi.org/10.1038/s41593-023-01563-6.
18. Nair, A., Karigo, T., Yang, B., Ganguli, S., Schnitzer, M. J., Linderman, S. W., Anderson, D. J., & Kennedy, A. An approximate line attractor in the hypothalamus encodes an aggressive state. Cell 186, 178-193.e15 (2023). https://doi.org/10.1016/j.cell.2022.11.027.
19. Wei, D., Osakada, T., Guo, Z., Yamaguchi, T., Varshneya, A., Yan, R., Jiang, Y., & Lin, D. A hypothalamic pathway that suppresses aggression toward superior opponents. Nat. Neurosci. 26, 774-787 (2023). https://doi.org/10.1038/s41593-023-01297-5.
20. Sofer, Y., Zilkha, N., Gimpel, E., Wagner, S., Chuartzman, S. G., & Kimchi, T. Sexually dimorphic oxytocin circuits drive intragroup social conflict and aggression in wild house mice. Nat. Neurosci. 27, 1565-1573 (2024). https://doi.org/10.1038/s41593-024-01685-5.
21. Allsop, S. A., Wichmann, R., Mills, F., Burgos-Robles, A., Chang, C. J., Felix-Ortiz, A. C., Vienne, A., Beyeler, A., Izadmehr, E. M., Glober, G., Cum, M. I., Stergiadou, J., Anandalingam, K. K., Farris, K., Namburi, P., Leppla, C. A., Weddington, J. C., Nieh, E. H., Smith, A. C., Ba, D., … Tye, K. M. Corticoamygdala transfer of socially derived information gates observational learning. Cell 173, 1329-1342.e18 (2018). https://doi.org/10.1016/j.cell.2018.04.004.
22. Jeon, D., Kim, S., Chetana, M., Jo, D., Ruley, H. E., Lin, S. Y., Rabah, D., Kinet, J. P., & Shin, H. S. Observational fear learning involves affective pain system and Cav1.2 Ca2+ channels in ACC. Nat. Neurosci. 13, 482-488 (2010). https://doi.org/10.1038/nn.2504.
23. Sterley, T. L., Baimoukhametova, D., Füzesi, T., Zurek, A. A., Daviu, N., Rasiah, N. P., Rosenegger, D., & Bains, J. S. Social transmission and buffering of synaptic changes after stress.

Nat. Neurosci. 21, 393-403 (2018). https://doi.org/10.1038/s41593-017-0044-6.

24. Qi, G., Zhang, P., Li, T., Li, M., Zhang, Q., He, F., Zhang, L., Cai, H., Lv, X., Qiao, H., Chen, X., Ming, J., & Tian, B. NAc-VTA circuit underlies emotional stress-induced anxiety-like behavior in the three-chamber vicarious social defeat stress mouse model. Nat. Commun. 13, 577 (2022). https://doi.org/10.1038/s41467-022-28190-2.

25. Yang, T., Bayless, D. W., Wei, Y., Landayan, D., Marcelo, I. M., Wang, Y., DeNardo, L. A., Luo, L., Druckmann, S., & Shah, N. M. Hypothalamic neurons that mirror aggression. Cell 186, 1195-1211.e19 (2023). https://doi.org/10.1016/j.cell.2023.01.022.

26. 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., Falkner, A. L., Shaevitz, J. W., … Murthy, M. SLEAP: A deep learning system for multi-animal pose tracking. Nat. Methods 19, 486-495 (2022). https://doi.org/10.1038/s41592-022-01426-1.

27. 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. Multi-animal pose estimation, identification and tracking with DeepLabCut. Nat. Methods 19, 496-504 (2022). https://doi.org/10.1038/s41592-022-01443-0.

28. Ghosh, K. K., Burns, L. D., Cocker, E. D., Nimmerjahn, A., Ziv, Y., Gamal, A. E., & Schnitzer, M. J. Miniaturized integration of a fluorescence microscope. Nat. Methods 8, 871-878 (2011). https://doi.org/10.1038/nmeth.1694.

29. Chen, T. W., Wardill, T. J., Sun, Y., Pulver, S. R., Renninger, S. L., Baohan, A., Schreiter, E. R., Kerr, R. A., Orger, M. B., Jayaraman, V., Looger, L. L., Svoboda, K., & Kim, D. S. Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature 499, 295-300 (2013). https://doi.org/10.1038/nature12354.

30. DeNardo, L. A., Liu, C. D., Allen, W. E., Adams, E. L., Friedmann, D., Fu, L., Guenthner, C. J., Tessier-Lavigne, M., & Luo, L. Temporal evolution of cortical ensembles promoting remote memory retrieval. Nat. Neurosci. 22, 460-469 (2019). https://doi.org/10.1038/s41593-018-0318-7.

31. Roth, B. L. DREADDs for neuroscientists. Neuron 89, 683-694 (2016). https://doi.org/10.1016/j.neuron.2016.01.040.

32. Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G. & Deisseroth, K. Millisecond-timescale,

genetically targeted optical control of neural activity. Nat. Neurosci. 8, 1263-1268 (2005). https://doi.org/10.1038/nn1525.

33. Govorunova, E. G., Sineshchekov, O. A., Janz, R., Liu, X. & Spudich, J. L. Natural light-gated anion channels: a family of microbial rhodopsins for advanced optogenetics. Science 349, 647-650 (2015). https://doi.org/10.1126/science.aaa7484.

34. Lebow, M. A. & Chen, A. Overshadowed by the amygdala: the bed nucleus of the stria terminalis emerges as key to psychiatric disorders. Mol. Psychiatry 21, 450-463 (2016). https://doi.org/10.1038/mp.2016.1.

35. Gungor, N. Z. & Paré, D. Functional heterogeneity in the bed nucleus of the stria terminalis. J. Neurosci. 36, 8038-8049 (2016). https://doi.org/10.1523/JNEUROSCI.0856-16.2016.

36. Adjei, M. P., Qasem, E., Aaflaq, S., Jacobs, J. T., Skinner, S., Summa, F., Spotanski, C., Thompson, R., Aholt, M. L., Lineberry, T., & Nordman, J. C. Familiarity gates socially transmitted aggression via the medial amygdala. J. Neurosci. 45, e1018252025 (2025). https://doi.org/10.1523/JNEUROSCI.1018-25.2025.

37. Iravedra-Garcia, J. M., Guthman, E. M., Cuturela, L., Ocasio-Arce, E. J., Pillow, J. W., & Falkner, A. L. Aggression experience and observation promote shared behavioral and neural changes. bioRxiv 2024.12.26.630396 (2024). https://doi.org/10.1101/2024.12.26.630396.

38. Mojahed, N., Adjei, M., Qasem, E., Aaflaq, S., Adu, T., Jacobs, J. T., Richardson, B. D., & Nordman, J. C. (2024). Acute social defeat during adolescence promotes long-lasting aggression through activation of the medial amygdala. Front. Neurosci. 18, 1433993 (2024). https://doi.org/10.3389/fnins.2024.1433993.

39. Newman, E. L., Covington, H. E., 3rd, Suh, J., Bicakci, M. B., Ressler, K. J., DeBold, J. F., & Miczek, K. A. Fighting females: neural and behavioral consequences of social defeat stress in female mice. Biol. Psychiatry 86, 657-668 (2019). https://doi.org/10.1016/j.biopsych.2019.05.005.

40. Whitten, C. J., Kearney, J. N., Radford, A. F., Hooker, M. K., Brewer, A. R., Kaminski, E. A., Kelly, J. R., Hennon, L. N., Perez, L. I., Jenkins, M. S., Schulz, K. M., & Cooper, M. A. Standing strong: effects of social dominance on neural activity and defeat-induced social avoidance in female Syrian hamsters. Horm. Behav. 177, 105868 (2026). https://doi.org/10.1016/j.yhbeh.2025.105868.

41. John, M. M., Pratt, M. A., Yaeger, J. D. W., Brummels, R. A., Ledesma, L. J., Meyer, L. S.,

Hartwig, R. L., Legner, G. L., Gilbertson, N. G., Ronan, P. J., & Summers, C. H. Aggression as a contributing factor to social defeat and stress vulnerability. Neurobiol. Stress 36, 100728 (2025). https://doi.org/10.1016/j.ynstr.2025.100728.

42. Nordman, J. C., Ma, X., Gu, Q., Potegal, M., Li, H., Kravitz, A. V., & Li, Z. Potentiation of divergent medial amygdala pathways drives experience-dependent aggression escalation. J. Neurosci. 40, 4858-4880 (2020). https://doi.org/10.1523/JNEUROSCI.0370-20.2020.

43. Nordman, J., Ma, X. & Li, Z. Traumatic stress induces prolonged aggression increase through synaptic potentiation in the medial amygdala circuits. eNeuro 7, ENEURO.0147-20.2020 (2020). https://doi.org/10.1523/ENEURO.0147-20.2020.

44. Zingg, B., Chou, X. L., Zhang, Z. G., Mesik, L., Liang, F., Tao, H. W., & Zhang, L. I. AAV-mediated anterograde transsynaptic tagging: mapping corticocollicular input-defined neural pathways for defense behaviors. Neuron 93, 33-47 (2017). https://doi.org/10.1016/j.neuron.2016.11.045.

45.Han, S., Ding, N., Li, C., & Yuan, P. In Vivo Calcium Imaging of Granule Cells in the Dentate Gyrus of Hippocampus in Mice. Journal of visualized experiments: JoVE, (210), 10.3791/66916 (2024). https://doi.org/10.3791/66916.

Supplemental figure legends

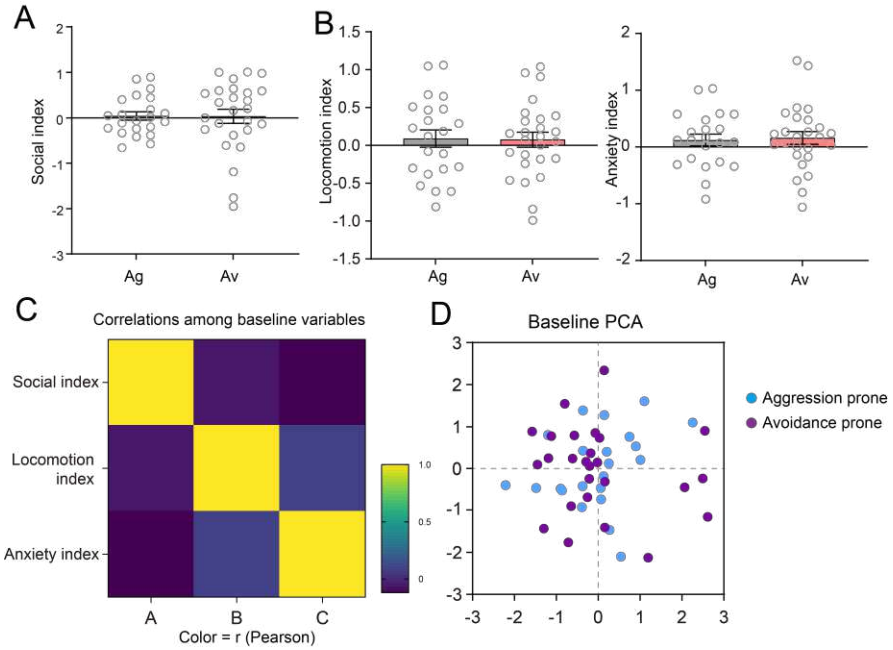


Figure S1. Baseline behavior does not pre-separate the later social states. (A) Baseline social index in mice subsequently assigned as aggression-prone (n = 22) or avoidance-prone (n = 26; p = 0.9847). (B) Baseline locomotion and anxiety indices in the same groups (p = 0.9172 and p = 0.8155). (C) Pearson correlations among baseline social, locomotion and anxiety indices. (D) Baseline PCA colored retrospectively by the later post-stress state, showing substantial overlap.

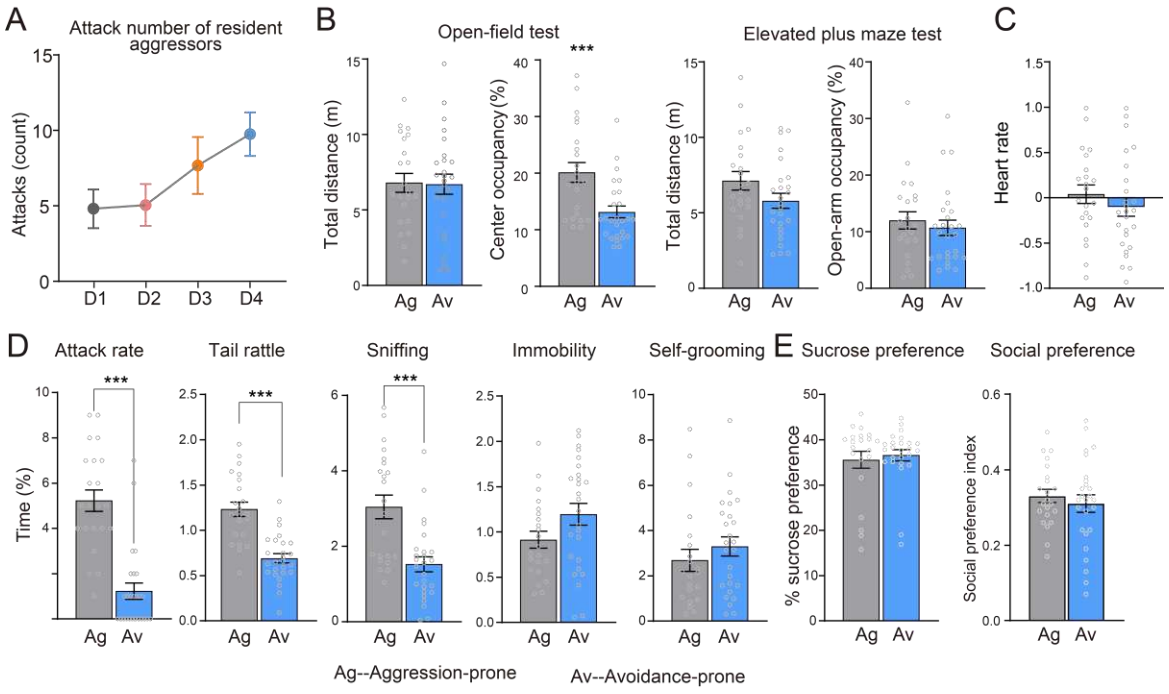


Figure S2. Behavioral validation and secondary phenotyping of post-stress states. (A) Resident-aggressor attack counts across screening sessions D1-D4 (n = 16). (B) Open-field distance (p = 0.922) and center occupancy (p = 0.001, t = 3.519), elevated-plus-maze distance (p = 0.0978) and open-arm occupancy (p = 0.5337) in aggression-prone and avoidance-prone observers. (C) Mouse-level heart-rate index. (D-E) Post-stress attack rate, tail-rattle duration, sniffing, immobility and self-grooming (n = 22 and 26), followed by sucrose and social preference (attack, $p < 0.0001$; tail rattling, $p < 0.0001$; sniffing, p = 0.0001; immobility, p = 0.0758; self-grooming, p = 0.3451; sucrose preference, p = 0.7381; social preference, p = 0.5749).

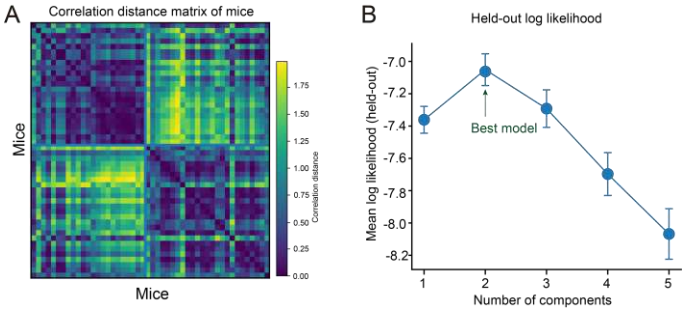


Figure S3. Robustness of behavioral-state inference. (A) PCA of the standardized five-feature post-stress matrix, colored only after unsupervised state assignment. (B) Held-out mean log likelihood across component number, estimated by repeated 5-fold cross-validation; K = 2 performed best (-7.05).

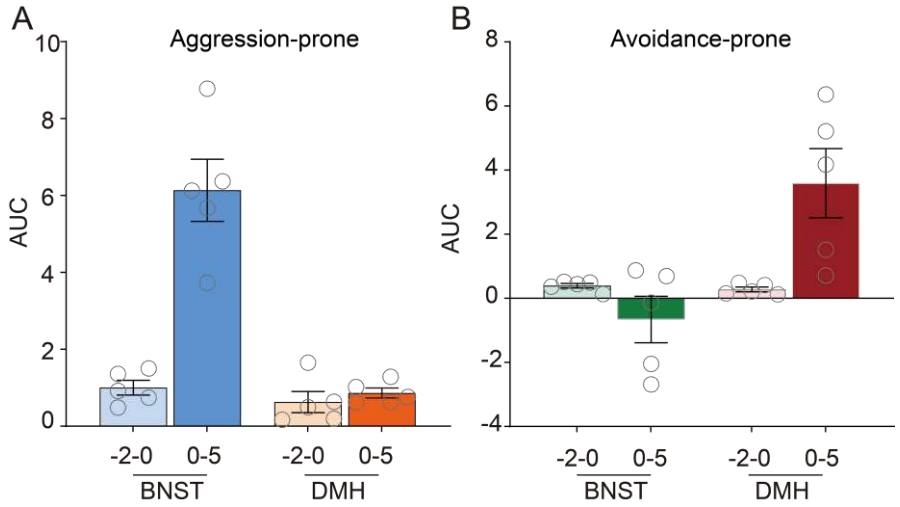


Figure S4. Complementary quantification of DP terminal-field bias using integrated activity. (A) Event-aligned BNST/DMH terminal AUC before and after the event in aggression-prone (p_{BNST} = 0.0019, t = 7.233; p_{DMH} = 0.4552, t = 0.826; n = 5 per state). (B) Event-aligned

BNST/DMH terminal AUC before and after the event in avoidance-prone mice (p_{BNST} = 0.2013, t = 1.528; p_{DMH} = 0.0315, t = 3.247; n = 5 per state).

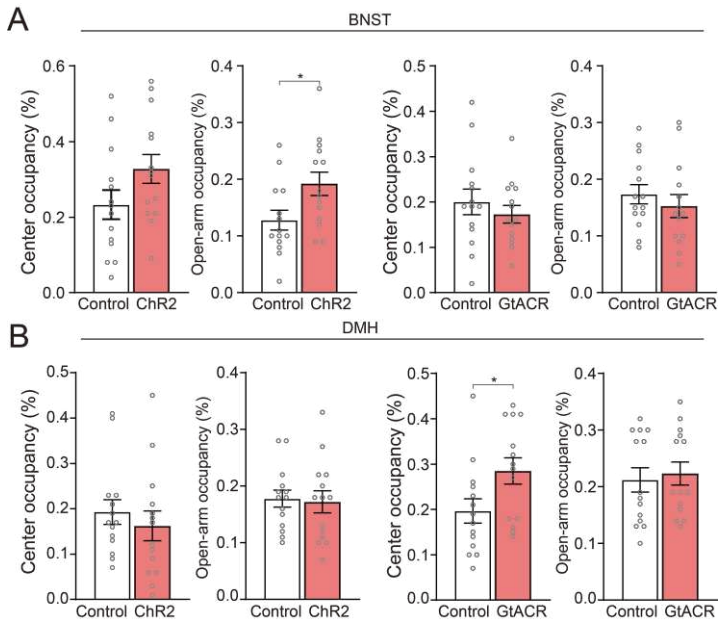


Figure S5. Nonsocial-anxiety controls for input-defined downstream manipulation. (A) Open-field center occupancy (p_{ChR2} = 0.0915, t = 1.752; p_{GtACR} = 0.4352, t = 0.7925) and elevated-plus-maze open-arm occupancy (p_{ChR2} = 0.0252, t = 2.376; p_{GtACR} = 0.4386, t = 0.7866) in BNST-targeted control and manipulation cohorts. (B) Open-field center occupancy (p_{ChR2} = 0.4792, t = 0.7179; p_{GtACR} = 0.033, t = 2.252) and elevated-plus-maze open-arm occupancy (p_{ChR2} = 0.8181, t = 0.2324; p_{GtACR} = 0.7007, t = 0.3887) in DMH-targeted control and manipulation cohorts.

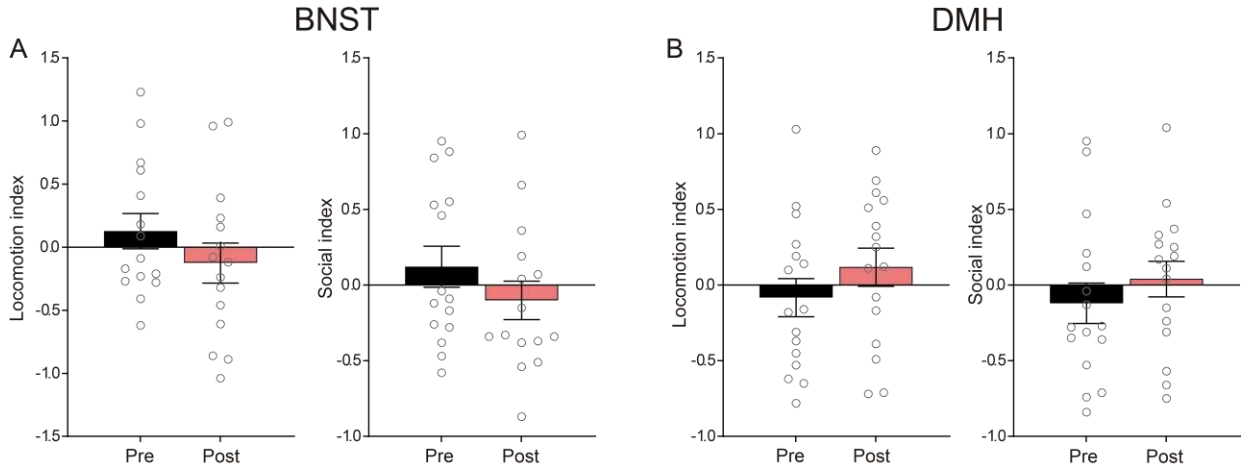


Figure S6. Nonsocial behavioral controls for cell-type-specific downstream manipulation. (A) Locomotor (p = 0.2834, t = 1.115) and social behavior index (p = 0.2583, t = 1.178) after chemogenetic inhibition of GABAergic BNST neurons in aggression-prone mice (n = 15). (B)

Locomotor (p = 0.3125, t = 1.045) and social behavior index (p = 0.3877, t = 0.8897) after chemogenetic activation of vGlut2-positive DMH neurons (n = 16).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarymaterial.zip