

Basal forebrain projections to the lateral habenula modulate aggression reward

Sam A. Golden^{1,2}, Mitra Heshmati^{1,2*}, Meghan Flanigan^{1,2*}, Daniel J. Christoffel³, Kevin Guise^{1,2}, Madeline L. Pfau^{1,2}, Hossein Aleyasin¹, Caroline Menard¹, Hongxing Zhang⁴, Georgia E. Hodes¹, Dana Bregman¹, Lena Khibnik¹, Jonathan Tai¹, Nicole Rebusi¹, Brian Krawitz^{1,2}, Dipesh Chaudhury⁴, Jessica J. Walsh³, Ming-Hu Han^{1,4}, Matt L. Shapiro¹ & Scott J. Russo¹

Maladaptive aggressive behaviour is associated with a number of neuropsychiatric disorders¹ and is thought to result partly from the inappropriate activation of brain reward systems in response to aggressive or violent social stimuli². Nuclei within the ventromedial hypothalamus^{3–5}, extended amygdala⁶ and limbic⁷ circuits are known to encode initiation of aggression; however, little is known about the neural mechanisms that directly modulate the motivational component of aggressive behaviour⁸. Here we established a mouse model to measure the valence of aggressive inter-male social interaction with a smaller subordinate intruder as reinforcement for the development of conditioned place preference (CPP). Aggressors develop a CPP, whereas non-aggressors develop a conditioned place aversion to the intruder-paired context. Furthermore, we identify a functional GABAergic projection from the basal forebrain (BF) to the lateral habenula (IHb) that bi-directionally controls the valence of aggressive interactions. Circuit-specific silencing of GABAergic BF–IHb terminals of aggressors with halorhodopsin (NpHR3.0) increases IHb neuronal firing and abolishes CPP to the intruder-paired context. Activation of GABAergic BF–IHb terminals of non-aggressors with channelrhodopsin (ChR2) decreases IHb neuronal firing and promotes CPP to the intruder-paired context. Finally, we show that altering inhibitory transmission at BF–IHb terminals does not control the initiation of aggressive behaviour. These results demonstrate that the BF–IHb circuit has a critical role in regulating the valence of inter-male aggressive behaviour and provide novel mechanistic insight into the neural circuits modulating aggression reward processing.

To study individual differences in aggression, we adapted the sensory contact model of social defeat for CD-1 mice^{9–11}, which exhibit a wide spectrum of aggressive behaviours. In this procedure (Fig. 1a), a sexually experienced adult male CD-1 mouse is presented with a series of novel 6–8-week-old subordinate male C57BL/6J intruder mice, who do not themselves exhibit any aggressive behaviours towards CD-1 mice (Extended Data Fig. 1a–i). This procedure identifies individual differences in antagonist aggressive behaviours without producing lasting stress-related behavioural phenotypes (Extended Data Table 1). Ethological analysis revealed that approximately 70% (310/448) of mice exhibited aggressive behaviour (termed aggressors (AGGs)) during at least one session, while approximately 30% (138/448) failed to initiate aggressive behaviour (termed non-aggressors (NONs)) at any time (Fig. 1b).

After repeated intruder interactions, AGGs have elevated serum testosterone (Fig. 1c) and decreased corticosterone (Fig. 1d) levels relative to NONs, suggesting that NONs may be less dominant and experience forced intruder interactions as more stressful. Analysis of several common metrics for aggression revealed normalized

distributions across AGGs that increased between screening sessions (Fig. 1e, f and Extended Data Fig. 2a–g). Importantly, the mean number of attack bouts (Extended Data Fig. 2f) and mean duration of attack bouts (Extended Data Fig. 2g) significantly correlate to mean attack latency. Therefore, attack latency provides a reliable index of aggression behaviours. Subsequently, we focused on AGGs that exhibited attack latencies within the most aggressive quartile of the sample distribution. These data confirm that outbred CD-1 mice exhibit a wide spectrum of aggressive behaviour and physiological responses to an intruder, leading us to hypothesize that there may be differences in the valence of intruder interactions among AGGs and NONs.

To assay the motivational state associated with intruder pairings, we developed an aggression-based CPP procedure. In this model, CD-1 mice are screened for aggression phenotype and then conditioned for CPP (Fig. 1g) by receiving novel C57BL/6J intruder-paired or intruder-unpaired sessions twice a day for three days. AGGs show a CPP for the intruder-paired context, while NONs show a conditioned place aversion (CPA) (Fig. 1h–j and Extended Data Fig. 3a–d). CPA in NONs does not appear to result from baseline differences in mood and anxiety or lack of interest in social targets (Extended Data Tables 1 and 2). However, we found that the valence of intruder interactions in AGGs and NONs is dependent upon intruder mice being freely moving and physically accessible during conditioning. Using a sensory CPP procedure in which the intruder mouse is placed in a protective cage within the intruder-paired context, both CPP and CPA are abolished (Fig. 1k–n and Extended Data Fig. 3e–h). These data demonstrate individual differences in the positive or negative valence of intruder interactions in AGGs versus NONs.

Clinical^{2,12} and preclinical⁸ studies have implicated BF structures, such as the nucleus accumbens (NAc), lateral septum and diagonal band nuclei (DBN), as potentially important brain regions controlling aggression-related behaviours. However, there has been limited functional evidence that the BF, or its projections, directly modulate the rewarding aspects of aggression. To define BF projections, we injected an adeno-associated virus (AAV) vector expressing enhanced yellow fluorescent protein (eYFP) under a neuronal-specific human synapsin (hSyn) promoter (AAV2-hSyn-eYFP) into the BF of CD-1 mice (Fig. 2a–c, top, and Extended Data Fig. 4a–c) targeted specifically to the more anterior septo-accumbal transition zone of the basal forebrain¹³ and observe a prominent axonal projection to the IHb (Fig. 2b, top).

To characterize BF–IHb projections further, we injected the IHb (Fig. 2a, b, bottom) with a retrograde monosynaptic glycoprotein-dead rabies virus (G-deleted-rabies-eGFP)¹⁴. Within the anterior BF that overlaps with our anterograde viral infection, we observed retrograde labelling in the septum (~45%), DBN (~35%) and the medial NAc shell (~15%) (Fig. 2c, d, bottom). Within retrogradely labelled BF slices, we performed *in situ* hybridization for GAD67, a marker of inhibitory

¹Fishberg Department of Neuroscience and Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA. ²Graduate Program in Neuroscience, Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA. ³Department of Psychiatry and Behavioral Sciences, Stanford University Medical Center, Palo Alto, California 94305, USA.

⁴Pharmacology and Systems Therapeutics and Institute for Systems Biomedicine, Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA.

*These authors contributed equally to this work.

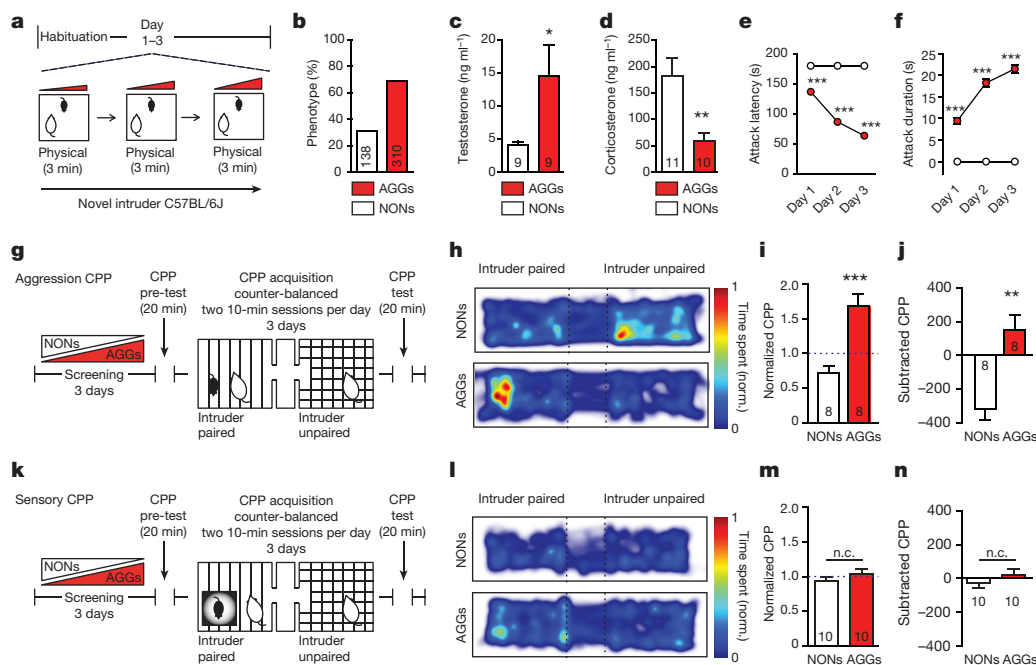


Figure 1 | Individual differences in aggression-related reward behaviour. **a**, Aggression screening: experimental schematic. **b**, Percentage of mice exhibiting aggressive (AGG) versus non-aggressive (NON) behaviours. **c**, **d**, Serum testosterone ($t_{16} = 2.23$, $*P < 0.05$; two-tailed unpaired t -test, $n = 9$ per group) (**c**) and corticosterone ($t_{19} = 3.231$, $**P < 0.01$; two-tailed unpaired t -test, $n = 10$ –11 per group) (**d**). **e**, **f**, Mean latency to attack (**e**) ($F_{2,1338} = 49.37$, two-way analysis of variance (ANOVA) $P < 0.001$; post-hoc test, $**P < 0.001$; $n = 138$ –310) and attack duration (**f**) ($F_{2,1338} = 22.35$, two-way ANOVA $P < 0.001$; post-hoc test, $**P < 0.001$; $n = 138$ –310). **g**, Aggression CPP schematic.

h, Representative heatmaps of aggression CPP. norm., normalized. **i**, **j**, Normalized ($t_{14} = 4.706$, $**P < 0.001$; two-tailed unpaired t -test, $n = 8$ per group) (**i**) and subtracted CPP score ($t_{14} = 4.013$, $**P < 0.01$; two-tailed unpaired t -test, $n = 8$ per group) (**j**). **k**, Sensory CPP schematic. **l**, Representative heatmaps of sensory CPP. **m**, **n**, Normalized (**m**) ($t_{18} = 1.023$, $P > 0.05$; two-tailed unpaired t -test, $n = 10$ per group) and subtracted (**n**) CPP score ($t_{18} = 0.961$, $P > 0.05$; two-tailed unpaired t -test, $n = 10$ per group). Summary data are represented as mean $\pm$ standard error of the mean. n.c., no change. Experiments were conducted once; n indicates biological replicates.

GABAergic neurons, and observed colocalization within the septum (~75%), DBN (~80%) and medial NAc shell (100%) (Fig. 2e).

To identify whether BF and lHb neurons are differentially activated by intruder interactions in AGGs and NONs, we examined c-Fos immunoreactivity 1 h after the final intruder screening. AGGs exhibit elevated c-Fos immunoreactivity in the septo-accumbal transition zone of the BF relative to NONs (Fig. 2f, g). Within the eYFP-positive BF terminal fields in the medial lHb, NONs exhibit increased c-Fos immunoreactive nuclei relative to AGGs (Fig. 2f, g). This finding was corroborated by slice electrophysiology, in which NONs exhibit an increase in lHb firing rates compared with AGGs 1 h after an intruder interaction that returns to baseline by 7 days after intruder interaction (Fig. 2h, i). Together, these data show that lHb neurons are differentially regulated by intruder interactions, possibly through inhibitory BF inputs.

To determine the functional contribution of BF–lHb projections, we conducted optogenetic circuit-specific terminal photostimulation in combination with slice electrophysiology with channelrhodopsin (AAV2-hSyn-ChR2(H134R)-eYFP) or halorhodopsin (AAV2-hSyn-NpHR3.0-eYFP), identifying photostimulation parameters that produce robust transient lHb activation or inhibition without rebound neuronal firing. ChR2^{BF→lHb} terminal photostimulation with 40 Hz resulted in significantly decreased lHb firing rates (Fig. 2j, k), while NpHR3^{BF→lHb} (8 s on, 2 s off) terminal photostimulation resulted in a robust increase in postsynaptic lHb firing rates (Fig. 2l, m). Importantly, whole-cell recordings from lHb neurons during ChR2^{BF→lHb} terminal photostimulation showed a significant increase in inhibitory postsynaptic currents (IPSCs) that was completely blocked by the GABA_A receptor antagonist, gabazine (Extended Data Fig. 4d, e). Optically induced IPSCs exhibited a response delay of ~7 ms (Extended Data Fig. 4f), which is in line with previously published response delays for ChR2 at monosynaptic circuits. Similarly, anterograde tracing of BF terminals in the lHb revealed that they were colocalized with vesicular

GABA transporter (VGAT), but not vesicular glutamate transporter 1 (VGLUT1) (Extended Data Fig. 4g). To validate these findings within an intact system, we used multi-electrode recording of postsynaptic lHb firing rates in anaesthetized mice in combination with terminal photostimulation (Extended Data Fig. 5a). Results show that activation (40 Hz ChR2^{BF→lHb}), or inhibition (8 s on, 2 s off NpHR3^{BF→lHb}) of presynaptic BF terminals in the lHb resulted in decreased or increased lHb postsynaptic neuronal firing, respectively (Extended Data Fig. 5b–d). These functional *in vitro* and *in vivo* recordings of ChR2^{BF→lHb} and NpHR3^{BF→lHb} confirm inhibitory GABAergic control over circuit activity and demonstrate reliable temporal control of lHb firing rates by optogenetic tools for *in vivo* behavioural analysis.

To investigate the functional consequences of BF–lHb neuronal firing on aggression reward, we paired photostimulation of ChR2^{BF→lHb} and NpHR3^{BF→lHb} in AGGs and NONs during the CPP test (Fig. 3a, b). NON::ChR2^{BF→lHb} stimulation promoted CPP (Fig. 3c–e), mimicking responses observed in control AGGs. Conversely, AGG::NpHR3^{BF→lHb} stimulation induced CPA (Fig. 3f–h), mimicking responses observed in control NONs. Neither NON::NpHR3^{BF→lHb} or AGG::ChR2^{BF→lHb} stimulation significantly affected the expression of CPP or CPA. Viral expression (Extended Data Fig. 6a–f) and locomotor activity (Extended Data Fig. 6g–j) were not different between conditions. These data confirm that BF–lHb circuitry modulates the rewarding component of aggressive behaviour and is both necessary and sufficient for the expression of CPP in AGGs and CPA in NONs.

To determine if these circuit-specific effects could be recapitulated by direct lHb cell body manipulation, we injected the lHb with AAV2-hSyn-ChR2-eYFP or AAV-hSyn-NpHR3.0-eYFP (Extended Data Fig. 7a–d) and directly stimulated lHb cell bodies using previously established optogenetic parameters for lHb¹⁵. NON::NpHR3.0^{lHb} stimulation to decrease lHb firing promoted CPP to the intruder-paired side (Extended Data Fig. 7e–g), whereas AGG::ChR2^{lHb} stimulation to

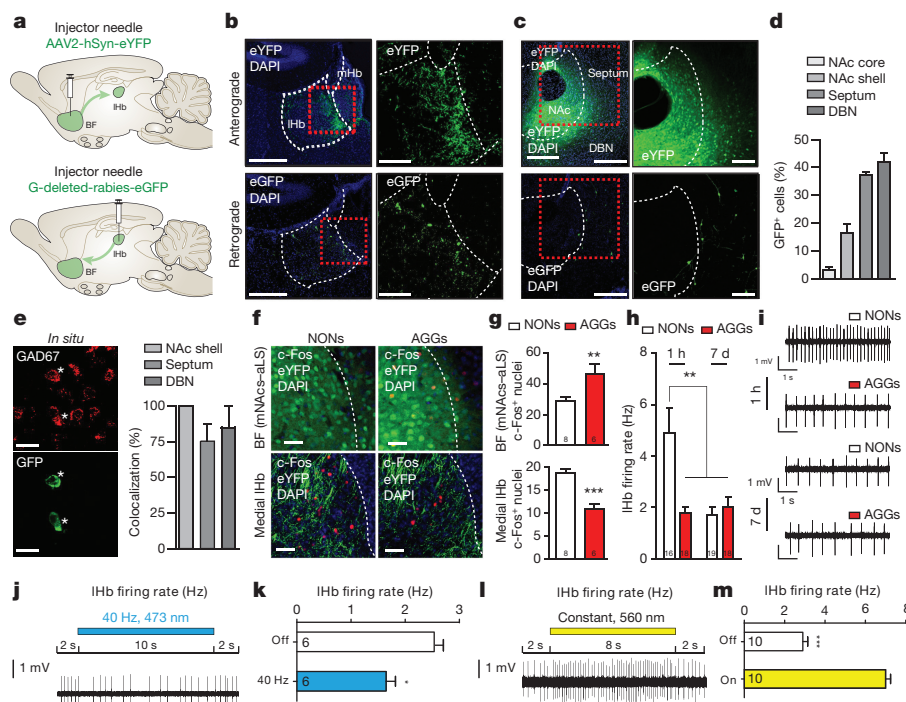


Figure 2 | GABAergic BF-IHb circuit is differentially activated by intruder interactions. **a**, Schematic of anterograde and retrograde tracing strategies. **b**, Representative anterograde AAV2-hSyn-eYFP infections (top, terminals) or retrograde G-deleted-rabies-eGFP infections (bottom, infection site) in IHb. Scale bars: 500 μ m; insets, 150 μ m. **c**, Representative anterograde AAV2-hSyn-eYFP infections (top, infection site) or retrograde G-deleted-rabies-eGFP infections (bottom, cell bodies) in the BF. Scale bars: 400 μ m; insets, 200 μ m. **d**, Percentage retrograde-labelled eGFP⁺ neurons within subnuclei of the anterior BF ($n = 3$ mice, ~229 cells per mouse). **e**, Representative *in situ* hybridization colocalized GAD67 and eGFP in DBN (left) and quantification (right) within the BF ($n = 3$ mice, 14 cells per mouse). Scale bars: 20 μ m. **f**, Representative images of AAV2-hSyn-eYFP infection and c-Fos immunoreactivity in medial NAc shell transition zone of the BF (top) and medial IHb terminals (bottom). Scale bars: 30 μ m. **g**, Quantification of c-Fos immunoreactivity in the medial NAc shell-septum transition zone ($t_8 = 2.655$, $*P < 0.05$; two-tailed

increase IHb firing induced CPA to the intruder-paired side (Extended Data Fig. 7h–j). Taken together, these results implicate the IHb as a key modulator of aggression motivational state.

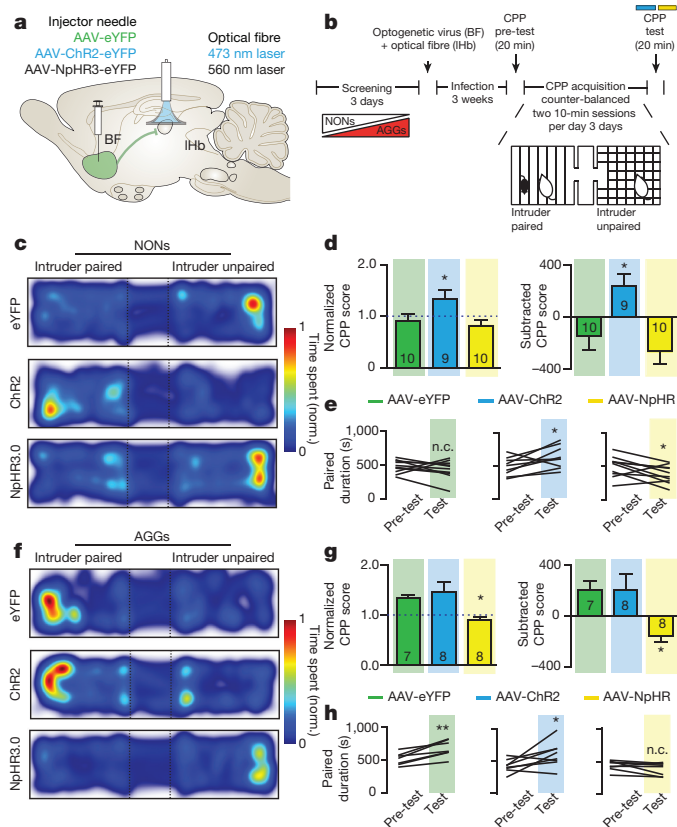
To determine if BF-IHb neuronal activity regulates the initiation or intensity of aggressive behaviour, we used ChR2^{BF→IHb} and NpHR3^{BF→IHb} (Fig. 4a) in AGGs and NONs during home-cage resident-intruder testing (Fig. 4b). Neither activation nor inhibition of BF-IHb terminals resulted in the initiation of aggressive behaviour (Fig. 4c, d), nor did it modulate social (Fig. 4e) and non-social (Fig. 4f) exploratory behaviours in NON mice. Similarly, AGG::ChR2^{BF→IHb} stimulation failed to initiate immediate attack behaviour, as indexed by no change in attack latency (Fig. 4g). However, AGG::ChR2^{BF→IHb} and AGG::NpHR3^{BF→IHb} stimulation bi-directionally modulated the severity of the aggressive behaviour relative to each other, although a nonsignificant trend was observed when either were compared to AGG::GFP^{BF→IHb} (Fig. 4h). As observed in NONs, AGG::ChR2^{BF→IHb} and AGG::NpHR3^{BF→IHb} photostimulation failed to alter either social (Fig. 4i) or non-social (Fig. 4j) exploratory behaviours. These data indicate that the BF-IHb circuit is important in modulating the intensity of aggressive behaviour; however, it is not a traditional attack initiation area.

On the basis of these data, we hypothesized that the BF-IHb circuit acts in other affective behavioural domains. We performed a behavioural battery to measure non-social generalized anxiety states and reward in naive CD-1 mice (Extended Data Fig. 8a). Both ChR2^{BF→IHb} and

unpaired t -test, $n = 6-8$ mice per group, 3 slices per mouse) and medial IHb ($t_{12} = 5.678$, $***P < 0.001$; two-tailed unpaired t -test, $n = 6-8$ mice per group, 3 slices per mouse). **h**, Firing rate of IHb neurons in AGG and NON mice at 1 h or 7 days after intruder interaction ($F_{1,67} = 10.56$, two-way ANOVA $P < 0.05$; post-hoc test, $**P < 0.01$; $n = 16-19$ cells per group, 4-5 mice per group). **i**, Representative trace of IHb *in vitro* cell-attached firing rates. **j**, Representative trace of IHb *in vitro* cell-attached firing rates during ChR2^{BF→IHb} photostimulation. **k**, Average firing rates of IHb neurons during ChR2^{BF→IHb} ($t_{10} = 3.679$, $**P < 0.01$; two-tailed unpaired t -test, $n = 6$ cells). **l**, Representative trace of IHb *in vitro* cell-attached firing rates during NpHR3^{BF→IHb} photostimulation. **m**, Average firing rates of IHb neurons during NpHR3^{BF→IHb} ($t_{18} = 11.68$, $***P < 0.0001$; two-tailed unpaired t -test, $n = 10$ cells) photostimulation. Data are represented as mean $\pm$ s.e.m. aLS, anterior lateral septum; DAPI, 4',6'-diamidino-2-phenylindole; mNAcs, medial nucleus accumbens shell. Experiments were conducted once; n indicates biological replicates.

NpHR3^{BF→IHb} terminal photostimulation failed to modulate anxiety-like behaviours in the open field (Extended Data Fig. 8b, c) and elevated plus maze tasks (Extended Data Fig. 8d, e). However, ChR2^{BF→IHb} stimulation potentiates the rewarding effects of cocaine by increasing the amount of time spent in the cocaine-paired chamber (Extended Data Fig. 8f). Therefore, while the BF-IHb circuit does not influence a generalized anxiety phenotype in the absence of social context or other stimuli, it does generalize to non-social rewarding stimuli such as cocaine.

Our results show individual differences in the rewarding properties of aggressive social interaction that are mediated by the BF-IHb circuit. When exposed to an intruder, AGGs exhibit increased activity of the BF and a concomitant reduction in IHb neuronal firing relative to NONs, contributing to a behavioural preference for environmental contexts associated with the interaction. Importantly, this circuit is not sufficient to induce attack behaviour. Although anatomical studies have identified this BF projection to the IHb in mice and rats^{16–18}, and diffusion tensor imaging (DTI) suggests probabilistic tract connections between the BF and IHb in humans¹⁹, this is the first study to provide functional evidence that GABAergic BF projections produce inhibitory control of IHb neurons to regulate the valence of aggressive intruder-based interactions. Stimulation or inhibition of BF-IHb projections is both sufficient and necessary to alter the positive or negative valence of an intruder-paired context. Our findings advance the understanding of IHb function in a behaviourally relevant animal model of aggression



motivation and provide further understanding into the physiology and neural circuitry of aggression and reward-related behaviours. While numerous functions have been ascribed to IHb²⁰ neuronal activity, including anxiety²¹, addiction²² and depression²³, there is a noticeable paucity of functional data addressing the role of IHb inputs, outside of those originating from the VTA region, within any of these behavioural domains. Indeed, anatomical tracing experiments have highlighted the complexity of IHb afferents²⁴ and efferents²⁵. With regard to the BF, the lateral septum, DBN and medial NAc shell, but not core, are known to send projections to the IHb^{24,26}. Our study implicates the septo-accumbal transition zone of BF as a critical source of GABAergic tone to the IHb within the context of motivated behaviour. However, on the basis of the fact that these BF GABAergic inputs to IHb

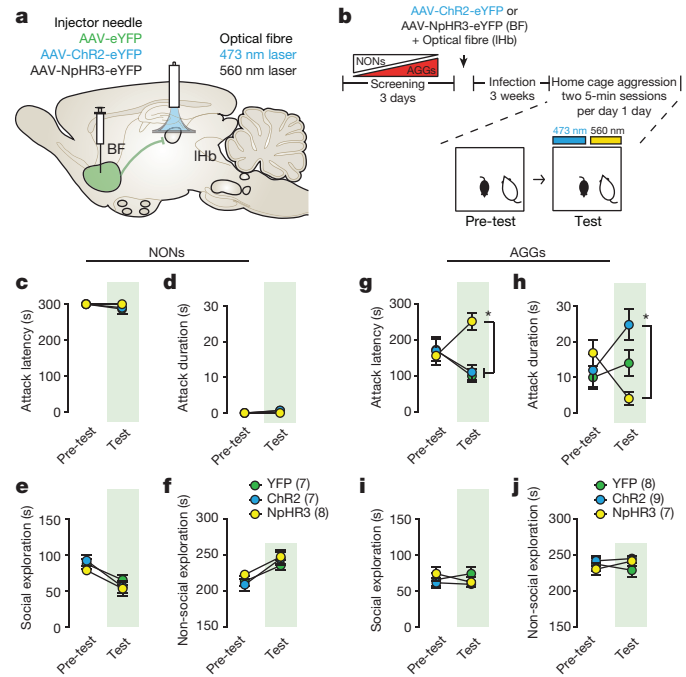


exhibit high tonic activation in acute slice that can be rapidly inhibited by terminal inhibition with NpHR3 (Fig. 2l, m), it is unlikely that they are NAc medium spiny neurons. Finally, based on both *in vitro* and *in vivo* electrophysiological studies, as well as anatomical tracing, we note that there may be a small subset of cells in the BF that either release an excitatory neurotransmitter or act indirectly on the IHb via di-synaptic inputs. It will be interesting in the future to determine what role these neurons have in reward processing.

Our results may provide important information to clinical studies identifying novel targets of deep brain stimulation in the treatment of neuropsychiatric conditions that present with aggression co-morbidity such as substance abuse²⁷ and depression²⁸. Deep brain stimulation protocols within specific BF nuclei²⁹ and the IHb³⁰ have been successfully used to treat intractable major depressive disorder, which is associated with symptoms of increased aggression in men²⁸. Overall, our findings demonstrate a previously unidentified functional role for the IHb and its inputs from the BF in mediating the rewarding component of aggression, and suggest that targeting shared underlying deficits in motivational circuitry may provide useful information for the development of novel therapeutic strategies for treating aggression-related neuropsychiatric disorders.

Our results may provide important information to clinical studies identifying novel targets of deep brain stimulation in the treatment of neuropsychiatric conditions that present with aggression co-morbidity such as substance abuse²⁷ and depression²⁸. Deep brain stimulation protocols within specific BF nuclei²⁹ and the IHb³⁰ have been successfully used to treat intractable major depressive disorder, which is associated with symptoms of increased aggression in men²⁸. Overall, our findings demonstrate a previously unidentified functional role for the IHb and its inputs from the BF in mediating the rewarding component of aggression, and suggest that targeting shared underlying deficits in motivational circuitry may provide useful information for the development of novel therapeutic strategies for treating aggression-related neuropsychiatric disorders.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 6 April 2015; accepted 17 May 2016.

- Anderson, D. J. Optogenetics, sex, and violence in the brain: implications for psychiatry. *Biol. Psychiatry* **71**, 1081–1089 (2012).
- Decety, J., Michalska, K. J., Akitsuki, Y. & Lahey, B. B. Atypical empathic responses in adolescents with aggressive conduct disorder: a functional MRI investigation. *Biol. Psychol.* **80**, 203–211 (2009).
- Yang, C. F. et al. Sexually dimorphic neurons in the ventromedial hypothalamus govern mating in both sexes and aggression in males. *Cell* **153**, 896–909 (2013).

4. Wasman, M. & Flynn, J. P. Directed attack elicited from hypothalamus. *Arch. Neurol.* **6**, 220–227 (1962).
5. Lin, D. *et al.* Functional identification of an aggression locus in the mouse hypothalamus. *Nature* **470**, 221–226 (2011).
6. Unger, E. K. *et al.* Medial amygdalar aromatase neurons regulate aggression in both sexes. *Cell Reports* **10**, 453–462 (2015).
7. Yu, Q. *et al.* Optogenetic stimulation of DAergic VTA neurons increases aggression. *Mol. Psychiatry* **19**, 635 (2014).
8. Takahashi, A. & Miczek, K. A. Neurogenetics of aggressive behavior: studies in rodents. *Curr. Top. Behav. Neurosci.* **17**, 3–44 (2013).
9. Kudryavtseva, N. N., Bakshantovskaya, I. V. & Koryakina, L. A. Social model of depression in mice of C57BL/6J strain. *Pharmacol. Biochem. Behav.* **38**, 315–320 (1991).
10. Golden, S. A., Covington, H. E. III, Berton, O. & Russo, S. J. A standardized protocol for repeated social defeat stress in mice. *Nat. Protocols* **6**, 1183–1191 (2011).
11. Miczek, K. A., DeBold, J. F. & Thompson, M. L. Pharmacological, hormonal, and behavioral manipulations in analysis of aggressive behavior. *Prog. Clin. Biol. Res.* **167**, 1–26 (1984).
12. Glenn, A. L. & Yang, Y. The potential role of the striatum in antisocial behavior and psychopathy. *Biol. Psychiatry* **72**, 817–822 (2012).
13. Zahm, D. S., Parsley, K. P., Schwartz, Z. M. & Cheng, A. Y. On lateral septum-like characteristics of outputs from the accumbal hedonic “hotspot” of Pecifia and Berridge with commentary on the transitional nature of basal forebrain “boundaries”. *J. Comp. Neurol.* **521**, 50–68 (2013).
14. Callaway, E. M. & Luo, L. Monosynaptic circuit tracing with glycoprotein-deleted rabies viruses. *J. Neurosci.* **35**, 8979–8985 (2015).
15. Lammel, S. *et al.* Input-specific control of reward and aversion in the ventral tegmental area. *Nature* **491**, 212–217 (2012).
16. Herkenham, M. & Nauta, W. J. Afferent connections of the habenular nuclei in the rat. A horseradish peroxidase study, with a note on the fiber-of-passage problem. *J. Comp. Neurol.* **173**, 123–145 (1977).
17. Sutherland, R. J. The dorsal diencephalic conduction system: a review of the anatomy and functions of the habenular complex. *Neurosci. Biobehav. Rev.* **6**, 1–13 (1982).
18. Lecca, S., Meze, F. J. & Mameli, M. The lateral habenula in addiction and depression: an anatomical, synaptic and behavioral overview. *Eur. J. Neurosci.* **39**, 1170–1178 (2014).
19. Shelton, L. *et al.* Mapping pain activation and connectivity of the human habenula. *J. Neurophysiol.* **107**, 2633–2648 (2012).
20. Hikosaka, O. The habenula: from stress evasion to value-based decision-making. *Nature Rev. Neurosci.* **11**, 503–513 (2010).
21. Lee, E. H. & Huang, S. L. Role of lateral habenula in the regulation of exploratory behavior and its relationship to stress in rats. *Behav. Brain Res.* **30**, 265–271 (1988).
22. Maroteaux, M. & Mameli, M. Cocaine evokes projection-specific synaptic plasticity of lateral habenula neurons. *J. Neurosci.* **32**, 12641–12646 (2012).
23. Li, B. *et al.* Synaptic potentiation onto habenula neurons in the learned helplessness model of depression. *Nature* **470**, 535–539 (2011).
24. Yetnikoff, L., Cheng, A. Y., Lavezzi, H. N., Parsley, K. P. & Zahm, D. S. Sources of input to the rostromedial tegmental nucleus, ventral tegmental area, and lateral habenula compared: a study in rat. *J. Comp. Neurol.* **523**, 2426–2456 (2015).
25. Quina, L. A. *et al.* Efferent pathways of the mouse lateral habenula. *J. Comp. Neurol.* **523**, 32–60 (2015).
26. Felton, T. M., Linton, L., Rosenblatt, J. S. & Morell, J. I. First and second order maternal behavior related afferents of the lateral habenula. *Neuroreport* **10**, 883–887 (1999).
27. Beck, A., Heinz, A. J. & Heinz, A. Translational clinical neuroscience perspectives on the cognitive and neurobiological mechanisms underlying alcohol-related aggression. *Curr. Top. Behav. Neurosci.* **17**, 443–474 (2014).
28. Martin, L. A., Neighbors, H. W. & Griffith, D. M. The experience of symptoms of depression in men vs women: analysis of the National Comorbidity Survey Replication. *JAMA Psychiatry* **70**, 1100–1106 (2013).
29. Bewernick, B. H., Kayser, S., Sturm, V. & Schlaepfer, T. E. Long-term effects of nucleus accumbens deep brain stimulation in treatment-resistant depression: evidence for sustained efficacy. *Neuropsychopharmacology* **37**, 1975–1985 (2012).
30. Sartorius, A. & Henn, F. A. Deep brain stimulation of the lateral habenula in treatment resistant major depression. *Med. Hypotheses* **69**, 1305–1308 (2007).

Acknowledgements This research was supported by US National Institutes of Health grants R01 MH090264, P50 MH096890 and P50 AT008661-01 (S.J.R.), R01 MH092306 (M.H.H.), T32 MH087004 (M.L.P., M.H. and M.F.), T32 MH096678 (M.L.P.), F30 MH100835 (M.H.), F31 MH105217 (M.L.P.), National Institute of General Medical Sciences 1F12GM117583-01 (S.A.G.) and the National Natural Science Foundation of China 81200862 (H.Z.). We would like to thank K. Miczek and Y. Shaham for their input.

Author Contributions S.A.G. and S.J.R. designed and wrote the manuscript. S.A.G., D.J.C., M.H., C.M., J.J.W., M.L.P., N.R. H.A., G.E.H., M.F., D.B., L.K., J.T. and B.K. collected behavioural and immunohistochemistry data and aided in data analysis. H.Z., M.-H.H., D.C., K.G. and M.L.S. designed, carried out and analysed electrophysiological experiments.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.J.R. (scott.russo@mssm.edu).

Reviewer Information *Nature* thanks O. Hikosaka and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Animals. Male CD-1 (ICR) mice (35–45 g, sexually experienced retired breeders; Charles River Laboratories (CRL)) were obtained at 4 months of age. All breeders were confirmed by CRL to have had equal access, experience and success as breeders. Male C57BL/6J mice (20–30 g; The Jackson Laboratory) were obtained at 7–8 weeks of age and used as novel intruders. All mice were allowed 1 week of acclimation to the housing facilities before the start of experiments. CD-1 mice were single housed, and C57BL/6J mice were group housed. All mice were maintained on a 12 h light:dark cycle with *ad libitum* access to food and water. Procedures were performed in accordance with the National Institutes of Health Guide for Care and Use of Laboratory Animals and the Icahn School of Medicine at Mount Sinai Institutional Animal Care and Use Committee.

Aggression screening and ethological analysis. Aggression screening was performed as previously described¹⁰. After a minimum of 1 week habituation to home cages, experimental CD-1 mice were exposed to a novel C57BL/6J intruder for 3 min daily over 3 consecutive days. Each intruder presentation was performed in the home cage of the CD-1 mouse between 1–3 PM daily under white light conditions. During screening sessions the cage top along with feeding and water apparatus were replaced with a clear Plexiglass cover to allow unimpeded viewing and video recording of screening sessions. The duration and number of screening sessions were selected to prevent induction of stress- and anxiety-related behaviours in CD-1 mice (Extended Data Tables 1 and 2), which has been shown to occur during extended antagonist encounters³¹. This allows for separation between aggression and stress-related states. All screening sessions were video recorded for later ethological analysis using a digital colour video camera. Ethological analysis of aggression behaviour was performed by two blinded observers, recording (1) latency to initial aggression, (2) the number of aggressive bouts, (3) the total duration of aggression, and (4) the mean duration of aggressive bouts. Operational definitions for these behaviours are defined as follows: initiation of aggression is defined by the first clear physical antagonist interaction initiated by the CD-1 mouse, not including grooming or pursuit behaviour. Aggressive bouts are defined by cycles of initiated aggression with continuous orientation by the CD-1 mouse towards the intruder, and only defined as completed when the CD-1 mouse has physically reoriented away from the intruder. This definition allows for slight breaks (less than 5 s) in continuous physical interaction within an aggressive bout, assuming the CD-1 mouse has remained oriented towards the intruder throughout. CD-1 mice were defined as AGG if they initiated aggression during any of the three screening sessions and NON were defined as those that showed no aggression during any screening sessions. All aggression screening was halted if an intruder showed any signs of injury in accordance with our previously published protocol¹⁰.

Aggression CPP and behavioural analysis. The aggression CPP protocol, developed on the basis of a previously published cocaine CPP protocol³², consisted of three phases: pre-test, acquisition, and test. Mice were acclimated to the testing facility for 1 h before testing. All phases were conducted under red light and sound-attenuated conditions. The CPP apparatus (Med Associates) consisted of two unique conditioning chambers with a neutral middle zone that allowed for unbiased entry into either conditioning chamber at the initiation of each trial. All CPP sessions were video recorded using Noldus Ethovision 3.0 (Noldus Information Technology). During the pre-test phase, mice were placed into the middle chamber of the conditioning apparatus and allowed to freely explore the full extent of the CPP apparatus for 20 min. There were no group differences in bias for either chamber, and conditioning groups were balanced in an unbiased fashion to account for start side preference. The acquisition phase consisted of three successive days with two conditioning trials each day for a total of six acquisition trials. Morning trials (between 8:00 AM and 10:00 AM) and afternoon trials (between 3:00 PM and 5:00 PM) consisted of CD-1 mice confined to one chamber for 10 min while in the presence or absence of a novel C57BL/6J intruder. All groups were counterbalanced for conditioning chamber. A total of three conditioning trials to the intruder-paired and intruder-unpaired context were performed. On the test day, CD-1 mice were placed into the middle arena of the CPP apparatus without an intruder and allowed to freely explore both chambers for 20 min. Analysis of duration spent within either context was used to identify a CPP or CPA to the intruder-paired context. For optogenetic experiments, stimulation was performed during the full duration of the test phase. Total locomotor responses were also assessed to ensure equal exploratory behaviour between groups. Behavioural analysis of aggression CPP data was performed by assessing (1) normalized CPP (test phase duration spent in the intruder-paired chamber divided by the pre-test duration spent in the intruder-paired chamber, accounting for behaviour during both sessions), (2) subtracted CPP (test phase duration spent in the intruder-paired chamber minus test phase duration spent in the intruder-unpaired chamber, accounting for test session behaviour only), and (3) group and individual durations in both pre-test and test sessions.

Sensory CPP. Sensory CPP was performed and analysed identically to the aggression CPP procedure, with the exception that the intruder C57BL/6J was placed within a physical barrier to provide only sensory contact with the resident CD-1 mice.

Cocaine CPP. A previously published cocaine CPP protocol³² was used, which consisted of three phases: pre-test, acquisition, and test. Mice were acclimated to the testing facility for 1 h before testing. All phases were conducted under red light and sound-attenuated conditions. The CPP apparatus (Med Associates) consisted of two unique conditioning chambers with a neutral middle zone that allowed for unbiased entry into either conditioning chamber at the initiation of each trial. All CPP sessions were video recorded using Noldus Ethovision 3.0 (Noldus Information Technology). During the pre-test phase, mice were placed into the middle chamber of the conditioning apparatus and allowed to freely explore the full extent of the CPP apparatus for 30 min. There were no group differences in bias for either chamber, and conditioning groups were balanced in an unbiased fashion to account for start side preference. The acquisition phase consisted of two successive days with two conditioning trials each day for a total of four acquisition trials. Morning trials (between 8:00 AM and 10:00 AM) and afternoon trials (between 3:00 PM and 5:00 PM) consisted of CD-1 mice confined to one chamber for 20 min paired with an intraperitoneal injection of cocaine (10 mg kg⁻¹); afternoon sessions were paired with saline injections. All groups were counterbalanced for conditioning chamber. On the test day, CD-1 mice were placed into the middle arena of the CPP apparatus and allowed to freely explore both chambers for 30 min. Analysis of duration spent within either context was used to identify a CPP or CPA to the cocaine-paired context. For optogenetic experiments, stimulation was performed during the full duration of the test phase. Total locomotor responses were also assessed to ensure equal exploratory behaviour between groups.

Sucrose preference. Sucrose preference was performed as previously described³³. One week after the final screening session, AGG and NON mice had their standard water bottle removed and replaced with two 50-ml conical tubes with sipper tops filled with water. After a 24-h habituation period, water from one 50-ml conical tube was replaced with 1% sucrose. All tubes were weighed, and mice were allowed 24 h to drink. Tubes were then reweighed, and their locations in the wire tops were switched before a second 24-h period of drinking. At the end of the second day of sucrose testing, preference was calculated as the total amount of sucrose consumption divided by the total amount of fluid consumed over the 2 days of sucrose availability.

Elevated plus maze. The elevated plus maze was performed as previously described³⁴. One week after the final screening session, AGG and NON mice were acclimated to the testing facility for 1 h before testing and then placed in the elevated plus maze under red light conditions for 5 min. Each arm of the maze measured 12 × 50 cm. The Plexiglas cross-shaped maze consisted of two open arms with no walls and two closed arms (40-cm-high walls) and was on a pedestal 1 m above floor level. Behaviour was tracked using an automated system (Noldus Ethovision; Noldus Interactive technologies). Behaviour was measured as total time in open and closed arms.

Open field and locomotor measures. Open field was performed as previously described³⁴. One week after the final screening session, AGG and NON mice were acclimated to the testing facility for 1 h before testing. Open-field testing was performed in black Plexiglas arenas (42 × 42 × 42 cm; Nationwide Plastics) under red light conditions. Testing sessions were 10 min long. Behaviour was tracked using an automated system (Noldus Ethovision; Noldus Interactive technologies) to record the total distance moved and time spent in the total arena and a delineated 'centre zone' (24 cm × 24 cm).

Forced-swim test. The forced-swim test was performed as previously described³⁴. One week after the final screening session, AGG and NON mice were placed in the test room for an hour before behavioural testing for habituation. Mice were tested in a 4-litre Pyrex glass beaker, containing 2 litres of water at 25 ± 1 °C for 6 min. Behaviour was videotaped (Noldus Ethovision; Noldus Interactive technologies) and analysed for duration immobile, duration mobile and total movement.

Social interaction (approach). Social approach testing was performed as previously described¹⁰. One week after the final screening session, AGG and NON mice were acclimated to the testing facility for 1 h before testing, and all testing was performed under red light conditions. Mice were placed in an open field black Plexiglas arena (42 × 42 × 42 cm; Nationwide Plastics) with a small animal cage placed at one end. Their movements were then automatically monitored and recorded (Ethovision 3.0; Noldus Information Technology) for 2.5 min in the absence (target absent phase) of a social target. This phase is used to determine baseline exploratory behaviour. We then immediately measured 2.5 min of exploratory behaviour in the presence of a caged novel CD-1 or C57BL/6J mouse (target present phase), again recording total distance travelled and duration of time spent in the interaction and corner zones. Social interaction behaviour is determined by the total time spent in each zone.

Novel object versus social target preference. The novel object versus social target test consisted of two phases: pre-test and test on consecutive days, as previously described³⁵. One week after the final screening session, AGG and NON mice were acclimated to the testing facility for 1 h before both sessions. All phases were run under red-light and sound-attenuated conditions. The testing apparatus (Med Associates) consisted of two identical chambers, with a neutral middle zone that allowed for unbiased entry into either chamber at the initiation of each trial. All sessions were video recorded from above (Noldus Ethovision 3.0, Noldus Information Technology) for later behavioural analysis. Briefly, during the pre-test phase, mice were placed into the middle chamber of the apparatus and allowed to freely explore all zones for 5 min. There were no group differences in pre-test preference for either chamber. Conditioning groups were then balanced in an unbiased way to account for individual animals' preference. On the test day, mice were placed back into the apparatus in the presence of both a novel object (an upside-down steel-bar pencil holder) on one side and a social target (identical pencil holder containing either a novel CD-1 or a C57BL/6J mouse) on the other. Test mice were allowed to freely explore the apparatus for 5 min. The time spent in each chamber was recorded and used for analysis. The subtracted social score is derived by subtracting time in social-paired chamber from time in novel object-paired chamber during the test phase. Normalized social score is the ratio of time spent in the chamber of interest (social target or novel object) during the test phase over the pre-test phase.

Blood sampling and testosterone/corticosterone ELISA. Submandibular vein bleeds were taken from mice 4–24 h after the final screening session as previously described³⁶. Serum testosterone (RND Systems, Testosterone Parameter Assay Kit) and corticosterone (Immunodiagnostic Systems, IDS Corticosterone EIA Kit) levels were assessed via ELISA according to manufacture specifications. Briefly, blood was collected in a serum separator tube, allowed to clot for 30 min at room temperature, and centrifuged at 1,000g for 15 min and stored frozen (−20°C) until analysis. The sensitivity of the testosterone assay (minimum detectable dose ranging from 0.012 to 0.041 ng ml^{−1}) falls well below the ranges detected experimentally within our cohort (lowest serum testosterone concentration of 2.58 ng ml^{−1}).

Perfusion and brain tissue processing. For immunohistochemistry and histology, mice were given a euthanizing dose of 15% chloral hydrate and transcardially perfused with cold 1% paraformaldehyde in PBS (pH 7.4) followed by fixation with cold 4% paraformaldehyde in PBS. Brains were dissected and post-fixed for 18 h in the same fixative. Coronal sections were prepared on a vibratome (Leica) at 50 µm to assess viral placement and immunohistochemistry. For *in situ* hybridization, mice brains were rapidly removed and flash frozen in −30°C isopentane for 60 s and then kept at −80°C until sectioning.

Immunohistochemistry, *in situ* hybridization and confocal microscopy. For c-Fos experiments, sections were incubated overnight in blocking solution (3% normal donkey serum, 0.3% Triton X-100 in PBS), washed in PBS for 2 h, then incubated for 48 h in primary antibody (rabbit anti-c-Fos (Santa Cruz Biotechnology, SC-42) 1:2,000). Slices were then washed in PBS for 2 h, incubated in secondary antibody for 2 h (donkey anti-rabbit Cy2 1:200 (Jackson ImmunoResearch)), then washed in PBS for 30 min before staining with 1 µg ml^{−1} DAPI (Sigma) for 20 min. Sections were then mounted, air-dried overnight and coverslipped with hardset Vectashield (Jackson ImmunoResearch). All slices were images using a Zeiss LSM 780. For c-Fos analysis, all images were taken at ×20 magnification for both the BF and IHB, using the tile-scan function to span the entire region of interest. Analysis of c-Fos-positive nuclei was performed using NIH ImageJ in conjunction with the 'analyze particle' function on single images. For representative images demonstrating the areas of viral infection, images were acquired at ×10 magnification using the tile-scan function.

For all other immunohistochemistry, coronal sections (50 µm) were used for all immunofluorescence experiments. Sections were incubated in blocking solution (3% normal donkey serum, 0.3% Triton X-100 in PBS) for 1 h. Sections were then incubated in primary antibody overnight at 4°C (VGAT 1:500 (Synaptic Systems); VGLUT1 (Millipore) 1:5,000; 1:250; GFP (Aves) 1:1,000). Next, sections were washed in PBS for 60 min and then incubated in secondary antibody for 2 h (donkey anti-guinea pig Cy5 1:400; donkey anti-goat Cy5 1:400; donkey anti-chicken Cy2 1:400 (Jackson ImmunoResearch)), then washed with PBS for 60 min, stained with 1 µg ml^{−1} DAPI (Sigma) for 20 min, mounted and air-dried overnight. Slides were quickly washed in ethanol 70%, 95%, 100% and Citrosolv (Fisher), and coverslipped with DPX mounting medium (Electron Microscopy Sciences). All slices were images using a Zeiss LSM 780. For puncta imaging, 1-µm z-stacks were taken at ×100 magnification for both the BF and IHB. Deconvolution was performed on all z-stacks with AutoQuant X (MediaCybernetics). For representative image demonstrating the area of viral infection, images were acquired at ×100 magnification using the tile-scan function.

For *in situ* hybridization, RNAscope Multiplex Fluorescent Kits (Advanced Cell Diagnostics) were used with the company-provided procedure. Briefly, fresh frozen brains were slide mounted at 16 µm thickness, fixed for 15 min in cold 4% PFA, seri-

ally dehydrated with increasing EtOH concentration washes (50%, 75% 100% EtOH for 2 min each), and pre-treated with protease reagent (Protease IV, RNAscope) for 20 min. Proprietary probes (Advanced Cell Diagnostics) for eGFP (Channel 1) or GAD67 (Channel 2) were hybridized at 40°C for 2 h, and then subjected to a series of amplification steps at 40°C (1-FL: 30 min; 2-FL: 15 min; 3-FL: 30 min; 4-FL: 15 min). For the fourth amplification step, Reagent Alt-A was used, corresponding with Channel 1 visualization at 488 nm and Channel 2 at 550 nm. Finally slides were treated for 2 min with DAPI, an immediately coverslipped with EcoMount.

***In vitro* electrophysiology.** All recordings were performed blind to experimental condition, and performed in both NON and AGG CD-1 mice. For optogenetic slice electrophysiology, mice were anaesthetized with isoflurane, and perfused with cold artificial cerebrospinal fluid (aCSF) composed of (in mM): 128 NaCl, 3 KCl, 1.25 NaH₂PO₄, 10 D-glucose, 24 NaHCO₃, 2 CaCl₂ and 2 MgCl₂ (oxygenated with 95% O₂ and 5% CO₂, pH 7.35, 295–305 mOsm) as described in our previous work^{37,38}. Briefly, acute brain slices containing the IHB were cut using a microslicer (DTK-1000, Ted Pella) in 95% O₂ and 5% CO₂ saturated sucrose-aCSF, which was derived by fully replacing NaCl with 254 mM sucrose. Slices were maintained in the holding chamber containing aCSF for 1 h at 37°C. Slices were then transferred into a recording chamber fitted with a constant flow rate of aCSF equilibrated with 95%/5% O₂/CO₂ (2.5 ml min^{−1}) maintained at 35°C. Cell-attached recording mode was used to measure the firing rates of IHB neurons. In these recording experiments, glass recording pipettes (7–10 MΩ) were filled with an internal solution composed of (in mM): 115 potassium gluconate, 20 KCl, 1.5 MgCl₂, 10 phosphocreatine, 10 HEPES, 2 magnesium ATP and 0.5 GTP (pH 7.2, 285 mOsm). For the experiments to measure inhibitory postsynaptic currents, whole-cell recordings were performed under voltage-clamp mode (holding at −70 mV) in the presence of kynurenic acid (1 mM) with or without gabazine (2 µM) in aCSF. Glass recording pipettes (3–4 MΩ) for these whole-cell studies were filled with the internal solution composed of (mM): 120 CsCl, 10 phosphocreatine-Na, 10 HEPES, 10 EGTA, 2 ATP-Mg, 0.3 GTP-Tris (pH 7.2, 285 mOsm). Data acquisition was conducted using a Digidata 1440A digitizer and pClamp 10.2 (Axon Instruments).

Stereotaxic surgery and viral gene transfer. All surgeries were performed under aseptic conditions using anaesthetic. Briefly, mice were anesthetized with a mixture of ketamine (100 mg per kg body weight) and xylazine (10 mg per kg body weight) and positioned in a small-animal stereotaxic instrument (David Kopf Instruments) and the skull surface was exposed. Thirty-three-gauge syringe needles (Hamilton Co.) were used to bilaterally infuse either 0.5 µl (BF) or 0.4 µl (IHB) of virus over a 5 min period and the needle was removed after 5 min. NAc shell–septum transition zone BF stereotaxic coordinates taken from bregma (anteroposterior +1.5 mm; mediolateral, +1.6 mm; dorsoventral, −4.4 mm; angle 10°). IHB stereotaxic coordinates taken from bregma (anteroposterior, −1.7 mm; mediolateral, +0.4 mm; dorsoventral, −2.5 mm; angle 0°). For IHB optogenetic experiments, animals were implanted with an optical fibre at the same time as viral injection (dorsoventral, −2.0 mm). For secure fixture of the implantable fibre to the skull, the skull was dried and then industrial-strength dental cement (Grip cement; Dentsply) was added between the base of the implantable fibre and the skull. For non-conditional axonal tract tracing, 0.5 µl AAV2-hSyn-eYFP (1.5 × 10^{11–13} infectious units per ml, UNC Vector Core) was injected bilaterally into the BF. For retrograde tracing, 0.4 µl G-deleted-rabies-eGFP (1.33 × 10⁸ infectious units per ml, Salk Gene Transfer Targeting and Therapeutics Core) was injected into the IHB. For behavioural optogenetic experiments, 0.5 µl of non-conditional AAV2-hSyn-eYFP, AAV2-hSyn-hChR2(H134R)-eYFP or AAV-hSyn-eNpHR3.0-eYFP (1.5 × 10^{11–13} infectious units per ml, UNC Vector Core) were injected into the BF (terminal stimulation) or IHB (cell body stimulation). All non-rabies AAV injections were performed between 4–6 weeks before tracing or behavioural experiments; rabies-infected brains were collected 7 days after injection.

Blue light stimulation. Optical fibres (Thor Labs, BFL37-200) were connected using an FC/PC adaptor to a 473-nm blue laser diode (Crystal Laser, BCL-473-050-M) and a stimulator (Agilent Technologies, no. 33220A) was used to generate blue light pulses. For all *in vivo* behavioural experiments and optrode recordings, mice were given 40 Hz 5 ms light stimulations. Intensity of light delivered to ferrule was ~10 mW. These parameters are consistent with previously validated and published protocols for NAc medium spiny neurons^{39,40} and IHB neurons^{15,41}.

Yellow light stimulation. Optical fibres (Thor Labs, SFS200/220Y) were connected using an FC/PC adaptor to a 561-nm yellow laser diode (Crystal Laser, CL561-050-L), and a stimulator (Agilent Technologies, no. 33220A) was used to generate yellow light pulses. For *in vivo* optrode recordings we tested a protocol of 8 s of yellow light on followed by 2 s of light off. Intensity of light delivered to ferrule was ~10 mW. These parameters are consistent with previously validated and published protocols for NAc MSNs and IHB neurons^{42,43}.

***In vivo* recordings**

Optrode construction and implantation. An optrode was constructed by gluing four tetrodes to an optical fibre. Four tetrodes spun of 12.7-µm-diameter nichrome wire

(California Fine Wire) were glued to a 200- μ m-diameter optical fibre (Thor Labs, SFS200/220Y) and cut so that they extended between 750 and 250 μ m beyond the end of the fibre. The tetrodes were pinned into an electrode interface board (EIB; Neuralynx) and the tips were plated by passing 0.2 μ A current pulses through the individual wires and a gold solution until the impedance reached 150–200 k Ω m. The optrode was mounted on a stereotax arm (Kopf Instruments) and then lowered into the brain during surgery. Two small holes were drilled anterior and posterior to the recording site to serve as sites for ground screws. The ground screws were constructed by soldering stainless steel self-tapping screws to 3 mm stainless steel wire secured to the EIB. Screws were inserted far enough to come in contact with dura.

Recording. Recordings were carried out using a Digital Lynx 16SX recording system and Cheetah data acquisition software (Neuralynx). Signals from the tetrodes were bandpass filtered between 600 and 9,000 Hz and digitized at 32 kHz. Spike detection was performed in real time using a thresholding procedure: when the filtered signal reached threshold amplitude on any wire, a sweep including 8 data points before the crossing and 24 points after (32 points, or 1 ms) were saved as a putative spike event. Spike sorting and noise filtering was performed offline. The laser intensity was adjusted to $\sim$ 5 mW at the tip of the optrode before implantation. The optrode was lowered using the stereotax arm until the tetrode tips reached the dorsal extent of the LHb. Once the tissue and recordings stabilized, the optrode was slowly advanced until spikes were observed on at least one of the tetrodes. Spike amplitude and firing rate were allowed to stabilize and observed for several minutes before recording. For all trials a 30 s baseline recording was acquired, followed by 1 min of stimulation and ending with a 30 s post-stimulation baseline. The optrode was then stepped forward and this procedure repeated until the inferior extent of the LHb was reached.

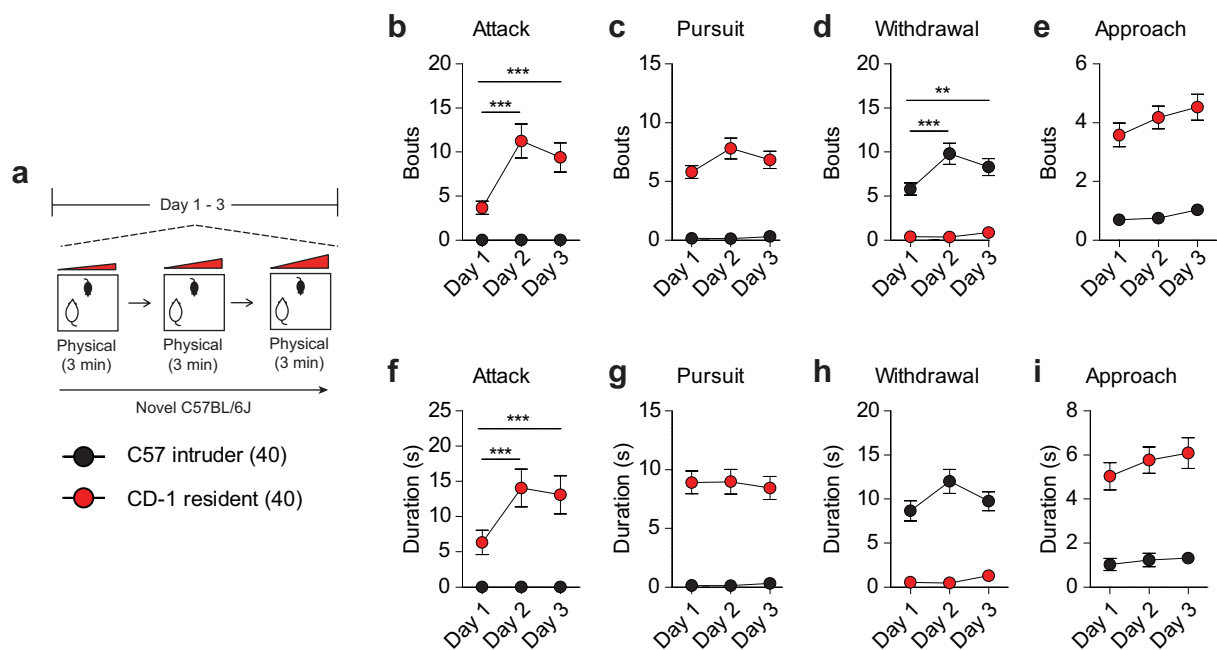
Analysis. Data were analysed using custom scripts written in Matlab (MathWorks). A first round of preliminary spike sorting was carried out using spike waveforms as parameters in KlustaKwik⁴⁴. The output from KlustaKwik was then imported into Matlab and clusters were manually edited using custom spike sorting software. Clearly separated clusters of spikes were assigned to functional units and entered into further analysis; noise spikes (for example, from spurious threshold crossings) and units that fired fewer than 100 spikes during recording were discarded. Spike rates were calculated in 2-s non-overlapping bins across the baseline and stimulation epochs. The resulting functions were smoothed using a Gaussian window with a standard deviation of 10 s. The rate function for each unit was then z-scored across all three epochs. For statistical analyses, rates were calculated in either 15-s bins or bins encompassing the entire baseline and stimulation periods. No smoothing was applied. The rate functions for each unit were z-scored across all three epochs and the z-scored rate functions were used to assess statistical significance.

Randomization and blinding. All experimenters were blinded to experimental condition. Mice were first screened to determine whether they were aggressive or

non-aggressive and then randomly assigned to optogenetic viral conditions for further behavioural analysis. For behavioural studies in Fig. 1 and slice physiology and c-Fos mapping studies in Fig. 2, AGGs and NONs were pre-screened for aggression and assigned to groups on the basis of their behavioural profile.

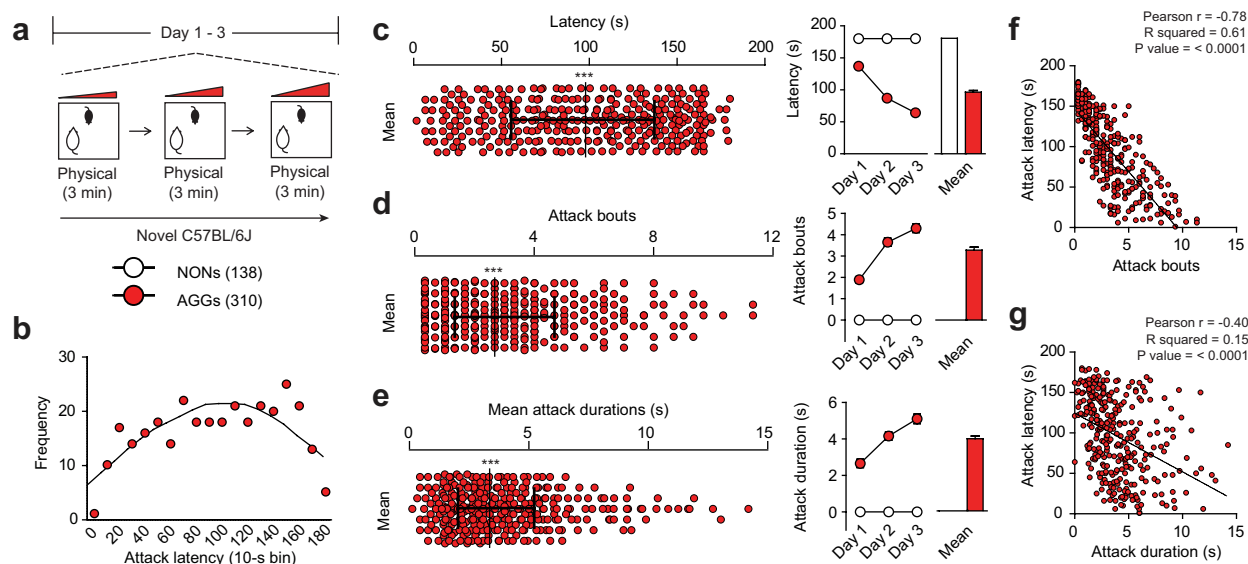
Statistical analysis. Sample size was calculated based on previous studies using Statmate from Graphpad prism (Graphpad Software). All *t*-tests, one-way ANOVAs, two-way ANOVAs and chi-squared tests were performed using Graph Pad Prism software (Graphpad Software Inc.). Bonferroni was used as a post-hoc test when appropriate for one-way and two-way ANOVAs. Normality was determined by D'Agostino–Pearson, Shapiro–Wilk and Kolmogorov–Smirnov normality tests. Statistical significance was set at $P < 0.05$.

31. Kudryavtseva, N. N., Bondar, N. P. & Avgustinovich, D. F. Association between experience of aggression and anxiety in male mice. *Behav. Brain Res.* **133**, 83–93 (2002).
32. Russo, S. J. *et al.* Nuclear factor kappa B signaling regulates neuronal morphology and cocaine reward. *J. Neurosci.* **29**, 3529–3537 (2009).
33. Golden, S. A. *et al.* Epigenetic regulation of RAC1 induces synaptic remodeling in stress disorders and depression. *Nature Med.* **19**, 337–344 (2013).
34. Krishnan, V. *et al.* Molecular adaptations underlying susceptibility and resistance to social defeat in brain reward regions. *Cell* **131**, 391–404 (2007).
35. Yang, M., Silverman, J. L. & Crawley, J. N. Automated three-chambered social approach task for mice. *Curr. Protoc. Neurosci.* Chapter 8, Unit 8 26 (2011).
36. Golde, W. T., Gollobin, P. & Rodriguez, L. L. A rapid, simple, and humane method for submandibular bleeding of mice using a lancet. *Lab Anim. (NY)* **34**, 39–43 (2005).
37. Chaudhury, D. *et al.* Rapid regulation of depression-related behaviours by control of midbrain dopamine neurons. *Nature* **493**, 532–536 (2013).
38. Friedman, A. K. *et al.* Enhancing depression mechanisms in midbrain dopamine neurons achieves homeostatic resilience. *Science* **344**, 313–319 (2014).
39. Lobo, M. K. *et al.* Δ FosB induction in striatal medium spiny neuron subtypes in response to chronic pharmacological, emotional, and optogenetic stimuli. *J. Neurosci.* **33**, 18381–18395 (2013).
40. Lobo, M. K. *et al.* Cell type-specific loss of BDNF signaling mimics optogenetic control of cocaine reward. *Science* **330**, 385–390 (2010).
41. Stamatakis, A. M. & Stuber, G. D. Activation of lateral habenula inputs to the ventral midbrain promotes behavioral avoidance. *Nature Neurosci.* **15**, 1105–1107 (2012).
42. Chandra, R. *et al.* Optogenetic inhibition of D1R containing nucleus accumbens neurons alters cocaine-mediated regulation of Tiam1. *Front. Mol. Neurosci.* **6**, 13 (2013).
43. Aquili, L., Liu, A. W., Shindou, M., Shindou, T. & Wickens, J. R. Behavioral flexibility is increased by optogenetic inhibition of neurons in the nucleus accumbens shell during specific time segments. *Learn. Mem.* **21**, 223–231 (2014).
44. Kadir, S. N., Goodman, D. F. M. & Harris, K. D. High-dimensional cluster analysis with the masked EM algorithm. *Neural Comput.* **26**, 2379–2394 (2014).



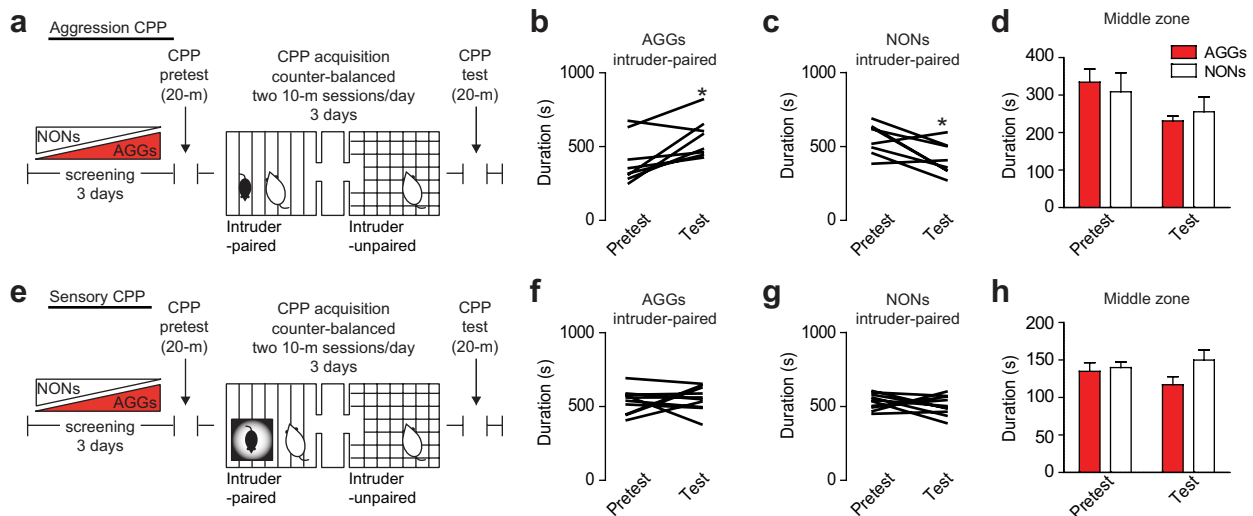
Extended Data Figure 1 | Social behaviours exhibited by resident CD-1 and intruder C57 mice during aggression screening. **a**, Experimental schematic of aggression screening procedure used in a subset (40 residents and 40 intruders) of mice to quantify social behaviours. **b–e**, Bouts of attacks ($F_{2,156} = 13.10$, two-way ANOVA $***P < 0.0001$; post-hoc test $***P < 0.001$; $n = 40$ per group) (b), pursuits (c), withdrawals

($F_{2,156} = 5.745$, two-way repeated measures ANOVA $***P < 0.001$; post-hoc test $***P < 0.001$; $n = 40$ per group) (d) and non-aggressive social approaches (e). **f–h**, Duration of attacks ($F_{2,156} = 7.069$, two-way repeated measures ANOVA $***P < 0.001$; post-hoc test $***P < 0.001$; $n = 40$ per group) (f), pursuits (g), withdrawals (h) and non-aggressive social approaches (e). All data are presented as mean $\pm$ s.e.m.



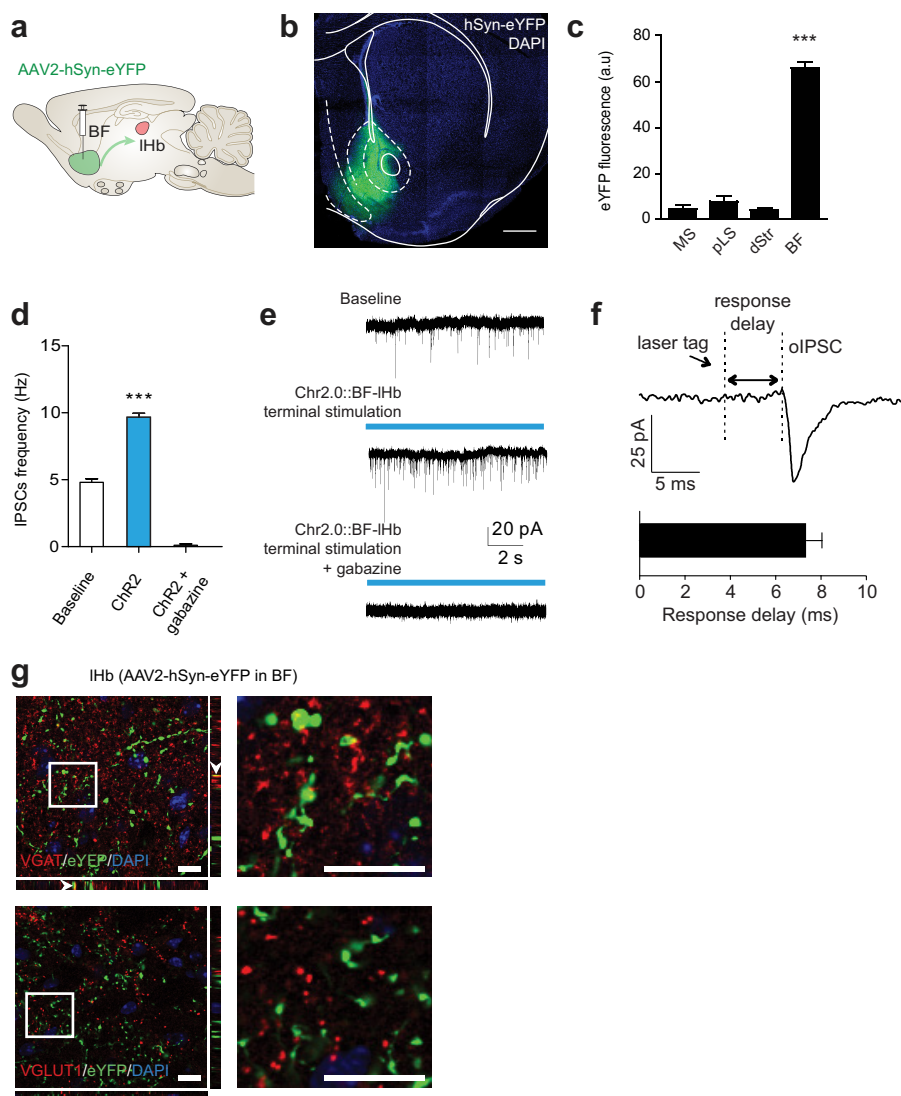
Extended Data Figure 2 | Detailed ethological analysis of AGG aggression-related behaviours. **a**, Experimental schematic of aggression screening procedure used in a sample (448 mice total; 138 NON and 310 AGG) of mice. **b**, Histogram of attack latency frequency using 10-s bins. **c–e**, Mean distribution across screening sessions (left) and individual screening sessions (right) for latency to aggression ($F_{2,1338} = 49.37$, two-way repeated measures ANOVA $P < 0.001$; post-hoc test, $*P < 0.001$; $n = 138–310$) (**c**), number of attack bouts ($F_{2,1338} = 21.03$, two-way repeated measures ANOVA $P < 0.001$; post-hoc test, $*P < 0.001$; $n = 138–310$) (**d**)

and mean attack duration ($F_{2,1338} = 11.96$, two-way repeated measures ANOVA $P < 0.001$; post-hoc test, $*P < 0.001$; $n = 138–310$) (**e**). **f**, **g**, Correlation of mean latency to initial aggression with mean attack bouts ($r = -0.78$, $P < 0.0001$) (**f**) and mean duration of attack bouts ($r = -0.40$, $P < 0.0001$) (**g**). Distribution plots are presented as the median with interquartile range and normality determined by D'Agostino–Pearson, Shapiro–Wilk and Kolmogorov–Smirnov normality tests ($P < 0.0001$). Summary data are represented as mean $\pm$ s.e.m.



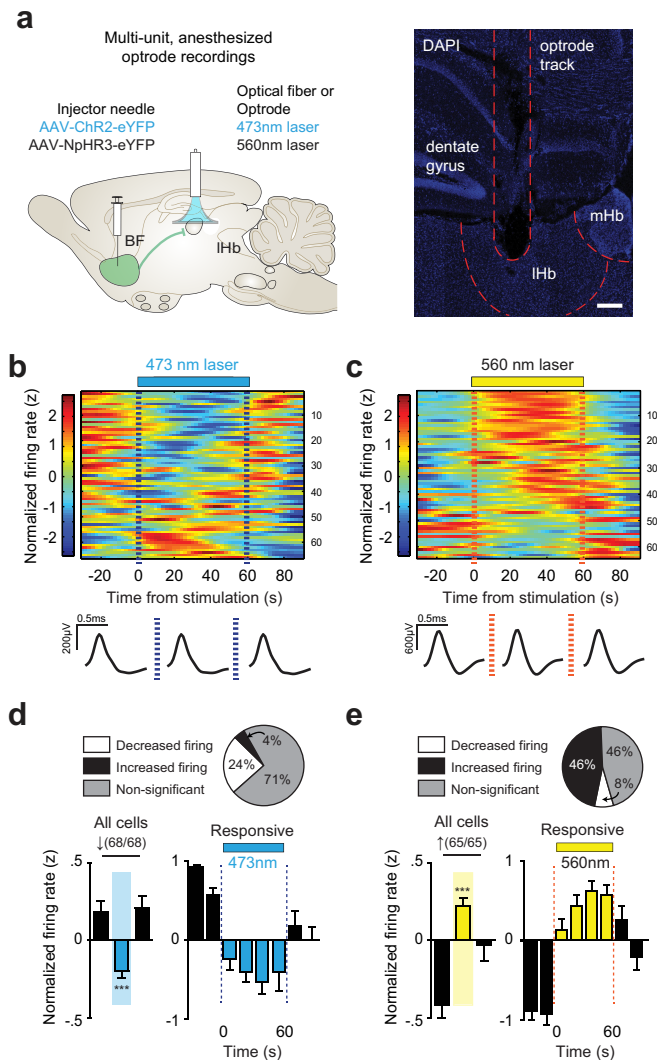
Extended Data Figure 3 | Aggression CPP behaviour. **a**, Experimental schematic of aggression CPP procedure. **b**, **c**, Individual duration spent in the intruder-paired context for AGG ($t_7 = 3.106$, $*P < 0.05$; two-tailed paired t -test, $n = 8$ per group) (**b**) and NON ($t_7 = 2.918$, $*P < 0.05$; two-tailed paired t -test, $n = 8$ per group) (**c**). **d**, Duration spent in the

middle neutral chamber during pre-test and test sessions. **e**, Experimental schematic of sensory CPP procedure. **f**, **g**, Individual duration spent in the intruder-paired context for AGG (**f**) and NON (**g**). **h**, Duration spent in the middle neutral chamber during pre-test and test sessions. Summary data are represented as mean $\pm$ s.e.m.



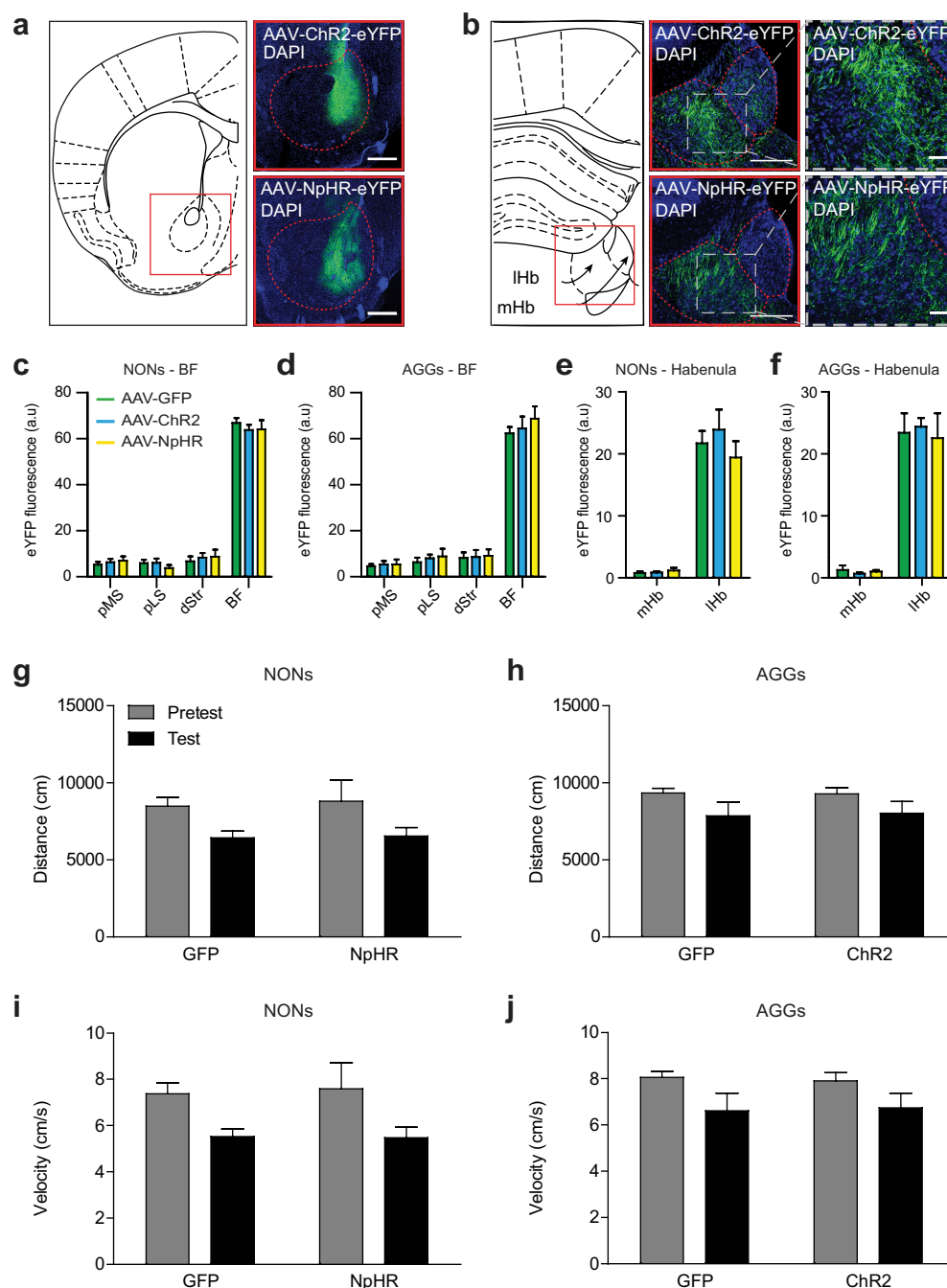
Extended Data Figure 4 | BF-IHb circuit tracing and GABAergic cell-type specificity. **a**, Schematic of viral tracing strategy. **b**, Representative BF viral infection with AAV2-hSyn-eYFP. Scale bar: 500 μ m. **c**, Histological analysis of viral infection with AAV2-hSyn-eYFP ($F_{3,11} = 223.0$, one-way ANOVA $***P < 0.0001$, post-hoc test, $***P < 0.0001$; $n = 3$ mice, 3 slices per mouse) across adjacent anatomical regions. **d**, **e**, Whole-cell electrophysiological recordings (**d**) and representative traces of IHb neurons photostimulated with AAV2-hSyn-Chr2.0 in the absence or

presence of bath-applied GABA_A receptor antagonist gabazine (2 μ m; $F_{2,7} = 220$, one-way ANOVA $P < 0.05$; post-hoc test, $***P < 0.001$, $n = 4$, 2, 2 cells from 2 mice) (**e**). **f**, Optically evoked IPSC response delay ($n = 21$ oIPSC events, 2 mice). **g**, Representative images of eYFP^{BF→IHb} terminal colocalization between vesicular GABA transporter (top), and not vesicular glutamate transporter 1 (bottom). Scale bars: 10 μ m; white arrows indicate colocalization within insets. MS, medial septum; pLS, posterior lateral septum. Summary data are represented as mean $\pm$ s.e.m.



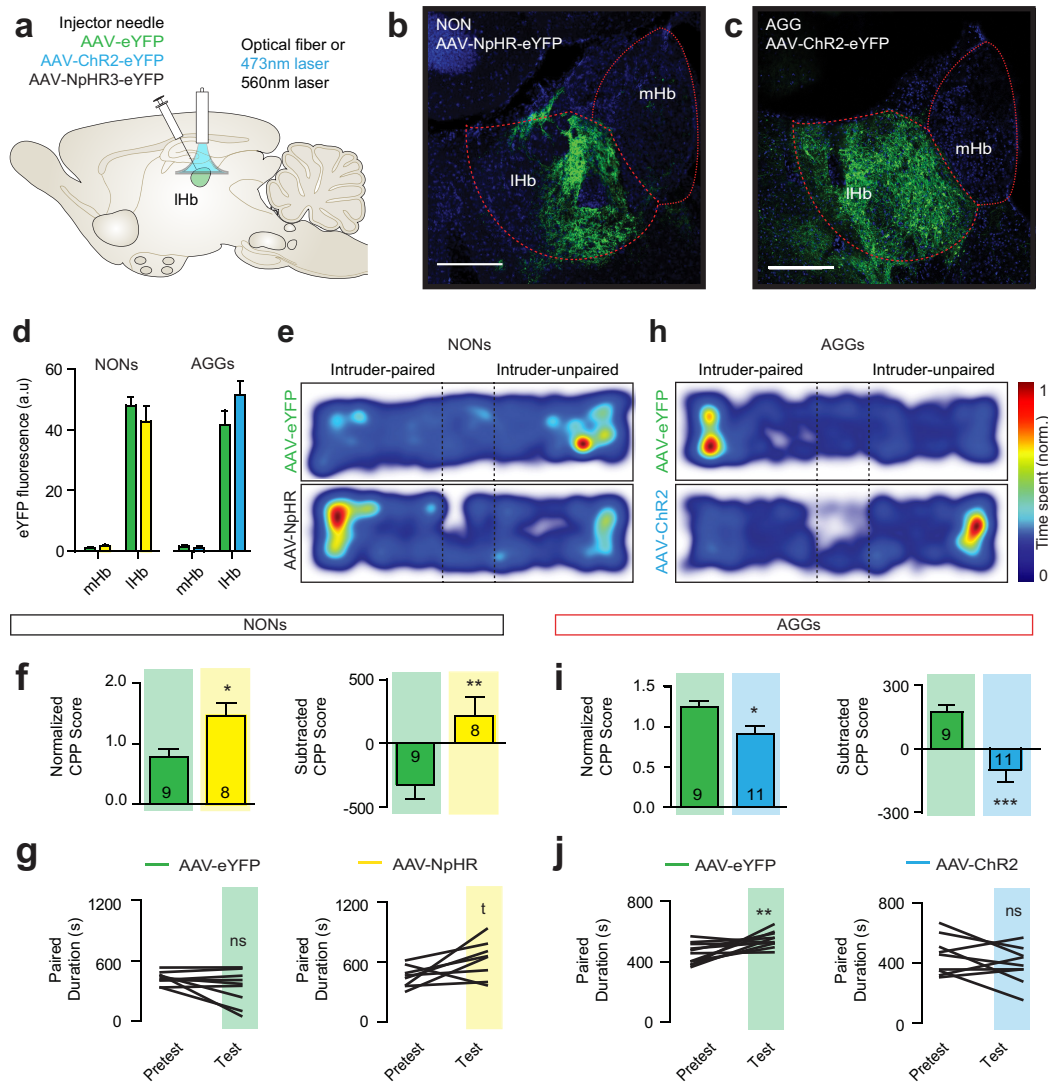
Extended Data Figure 5 | Multiunit anaesthetized optrode recordings.

a, Schematic of *in vivo* anaesthetized multi-unit optrode recording procedure (left) and representative optrode placement in LHb (right; scale bar: 200µm). **b**, **c**, Heatmaps of normalized firing rates for LHb neurons in response to BF terminal stimulation with ChR2^{BF→LHb} (**b**) or NpHR3^{BF→LHb} (**c**) and averaged spike wave-form shown below for pre-stimulation, stimulation and post-stimulation epochs. **d**, Percentage of cells by firing response (top) and average normalized LHb firing rate (bottom) after BF–LHb terminal stimulation with ChR2^{BF→LHb} for all identified cells ($F_{2,134} = 8.249$, one-way repeated-measure ANOVA $P < 0.001$; post-hoc test, $*P < 0.05$; $n = 68$ cells from 3 mice) and cells that significantly decreased firing during the stimulation epoch ($F_{7,105} = 8.868$, one-way repeated-measure ANOVA $P < 0.0001$; post-hoc test, $*P < 0.05$; $n = 16/68$ cells from 3 mice). **e**, Percentage of cells by firing response (top) and average normalized LHb firing rate (bottom) after BF–LHb terminal stimulation with NpHR3^{BF→LHb} for all identified cells ($F_{2,128} = 10.32$, one-way repeated-measure ANOVA $P < 0.0001$; post-hoc test, $*P < 0.05$; $n = 65/65$ cells from 3 mice) and cells that significantly increased firing during the stimulation epoch ($F_{7,203} = 17.58$, one-way repeated-measure ANOVA $P < 0.0001$; post-hoc test, $*P < 0.05$; $n = 30/65$ cells from 3 mice). mHb, medial habenula. Summary data are represented as mean $\pm$ s.e.m.



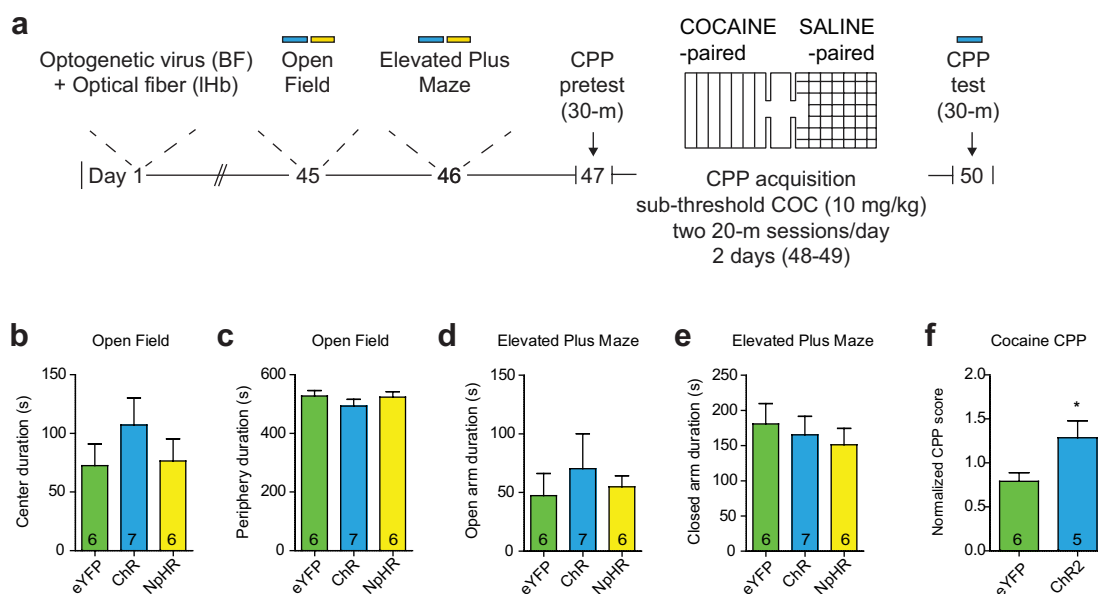
Extended Data Figure 6 | BF-IHb AAV infection and CPP locomotor behaviour. **a**, Schematic of BF coronal slice (left), alongside representative AAV-ChR2-eYFP (top) and AAV-NpHR3.0-eYFP (bottom) infections. Scale bar: 500 μ m. **b**, Schematic of IHb coronal slice (left), alongside representative images of BF terminal infection by AAV-ChR2-eYFP (middle top) and AAV-NpHR3.0-eYFP (middle bottom) within the IHb. Scale bar: 200 μ m. Representative close-ups of terminal regions shown in insets on right. Scale bar: 50 μ m. All representative images counterstained

with DAPI. **c**, **d**, Histological analysis of BF infection in NON (**c**) and AGG (**d**) mice. **e**, **f**, Histological analysis of habenular viral infection in NON (**e**) and AGG mice (**f**). **g**–**j**, Total distance travelled (**g**, **h**) and mean velocity (**i**, **j**) between NON and AGG during the CPP pre-test and test phase. All data are presented as mean $\pm$ s.e.m., and are not significant as determined by two-way ANOVA, $P < 0.05$. dStr, dorsal striatum; mHb, medial habenula; MS, medial septum; pLS, posterior lateral septum.



Extended Data Figure 7 | Direct IHb stimulation bi-directionally modulates aggression reward. **a**, Schematic of viral infection strategy. **b**, **c**, Representative images of IHb cell body infection in NON (**b**) and AGG (**c**). Scale bar: 200 μ m. **d**, Histological analysis of IHb viral infection. **e**, Representative CPP traces of NON. NON::NpHR^{IHb} cell body infection mimics the physiological effect of NON::ChR2^{BF→IHb} terminal stimulation. **f**, Normalized CPP score ($t_{15} = 2.834$, $*P < 0.05$; two-tailed unpaired t -test, $n = 8-9$ per group) and subtracted CPP score ($t_{15} = 3.058$, $**P < 0.01$; two-tailed unpaired t -test, $n = 8-9$ per group) in NON::eYFP^{IHb} and NON::NpHR^{IHb}. **g**, Individual duration spent in the intruder-paired context for NON::eYFP^{IHb} ($t_9 = 0.9129$, $P > 0.05$; two-tailed

paired t -test, $n = 10$ per group) and NON::NpHR^{IHb} ($t_9 = 2.344$, $*P < 0.05$; two-tailed paired t -test, $n = 10$ per group). **h**, Representative CPP traces of AGG::eYFP^{IHb} and AGG::ChR2^{IHb}. **i**, Normalized CPP score ($t_{18} = 2.692$, $*P < 0.05$; two-tailed unpaired t -test, $n = 9-11$ per group) and subtracted CPP score ($t_{18} = 4.203$, $***P < 0.01$; two-tailed unpaired t -test, $n = 9-11$ per group) for the intruder-paired context in AGG::eYFP^{IHb} and AGG::ChR2^{IHb}. **j**, Individual duration spent in the intruder-paired context for AGG::eYFP^{IHb} mice ($t_{10} = 3.212$, $**P < 0.01$; two-tailed paired t -test, $n = 9$ per group) and AGG::ChR2^{IHb} mice ($t_8 = 1.348$, $P < 0.05$; two-tailed paired t -test, $n = 11$ per group). Summary data are represented as mean $\pm$ s.e.m. dStr, dorsal striatum; mHb, medial habenula.



Extended Data Figure 8 | BF-IHb stimulation modulates cocaine CPP. **a**, Experimental timeline of general anxiety and cocaine CPP testing. **b–e**, BF-IHb stimulation during open field testing (**b**, **c**) and elevated plus maze testing (**d**, **e**). **f**, Subthreshold cocaine (10 mg kg⁻¹, intraperitoneal) CPP procedure with BF-IHb stimulation during CPP test ($t_9 = 2.403$, $P < 0.05$; two-tailed unpaired t -test, $n = 5–6$ per group).

Extended Data Table 1 | Stress and anxiety behaviours in AGG and NON

Elevated plus maze

		Duration in (s)		Latency to enter (s)		Distance (cm)	Velocity (cm/s)
		Closed arms	Open arms	Closed arms	Open arms		
NON	Mean	136.53	58.85	13.51	9.06	2412.14	8.16
	SEM	11.71	8.55	9.22	2.37	114.19	0.38
AGG	Mean	137.39	65.07	5.11	5.50	2515.77	8.67
	SEM	6.51	10.48	2.21	1.79	136.55	0.46
P value		0.95	0.65	0.39	0.24	0.57	0.40
n		12	12	12	12	12	12

Open field and locomotion

		Duration in (s)			Latency (s)	Distance (cm)	Velocity (cm/s)
		Center	Middle	Periphery	Center		
NON	Mean	18.80	126.63	373.76	24.35	4909.51	8.77
	SEM	3.28	28.87	58.22	7.78	374.32	0.59
AGG	Mean	15.59	138.70	445.82	31.31	4335.25	8.90
	SEM	1.77	11.63	12.20	7.22	211.61	0.36
P value		0.36	0.82	0.48	0.53	0.73	0.84
n		7-11	7-11	7-11	7-11	7-11	7-11

Forced-swim test

		Duration (s)		Distance (cm)	Velocity (cm/s)	Percent immobile
		Mobile	Immobile			
NON	Mean	260.96	175.76	2041.17	11.40	0.40
	SEM	7.78	8.21	101.09	0.58	0.02
AGG	Mean	275.40	164.19	2076.95	11.57	0.37
	SEM	8.20	8.92	70.68	0.39	0.02
P value		0.22	0.35	0.78	0.81	0.29
n		10	10	10	10	10

Sucrose preference

		Volume consumed (ml)						Sucrose preference		
		Day 1		Day 2		Day 1 + Day 2		Sucrose / (Sucrose + Water)		
		Water	Sucrose	Water	Sucrose	Water	Sucrose	Day 1	Day 2	Day 1 + Day 2
NON	Mean	2.44	5.74	1.54	5.74	3.98	11.48	0.71	0.76	0.75
	SEM	0.65	0.66	0.26	0.69	0.79	0.86	0.07	0.05	0.04
AGG	Mean	2.39	5.15	2.03	7.50	4.42	12.65	0.69	0.78	0.74
	SEM	0.58	0.58	0.59	0.77	0.70	1.05	0.06	0.06	0.04
P value		0.95	0.51	0.47	0.11	0.68	0.41	0.82	0.77	0.84
n		10-11	10-11	10-11	10-11	10-11	10-11	10-11	10-11	10-11

The behavioural data are shown as mean $\pm$ s.e.m. and analysed by unpaired Student's *t*-test. Significance at **P* < 0.05.

Extended Data Table 2 | Social approach behaviours in AGG and NON

Social interaction test

<u>Approach</u>		No target interaction zone (s)		Target interaction zone (s)		No target corner zone (s)		Target corner zone (s)	
		CD-1	C57	CD-1	C57	CD-1	C57	CD-1	C57
NON	Mean	56.62	51.55	77.49	76.42	34.77	32.53	21.54	25.03
	SEM	7.20	7.66	5.25	6.71	5.64	6.23	4.76	2.80
AGG	Mean	55.71	54.17	82.44	79.37	28.24	29.83	20.79	23.20
	SEM	4.67	6.07	6.13	8.91	2.52	5.71	2.84	4.95
P value		0.67	0.80	0.53	0.85	0.56	0.55	0.81	0.82
n		6-7	6-7	6-7	6-7	6-7	6-7	6-7	6-7

<u>Locomotion</u>		No target distance (cm)		Target distance (cm)		No target velocity (cm/s)		Target velocity (cm/s)	
		CD-1	C57	CD-1	C57	CD-1	C57	CD-1	C57
NON	Mean	1501.64	1715.50	1343.64	1447.93	10.01	11.44	8.96	9.65
	SEM	198.09	83.30	153.22	89.24	1.32	0.56	1.02	0.59
AGG	Mean	1544.73	1699.40	1284.39	1474.84	10.31	11.39	8.57	9.87
	SEM	31.04	121.71	41.37	97.81	0.20	0.83	0.28	0.63
P value		0.83	0.92	0.72	0.84	0.82	0.96	0.72	0.81
n		6-7	6-7	6-7	6-7	6-7	6-7	6-7	6-7

Social vs. novel object test

<u>Pretest phase</u>		Duration (s)		CD-1 Target Normalized social score		Subtracted social score	Duration (s)		C57BL6/J Target Normalized social score		Subtracted social score
		Social	Novel	Social	Novel		Social	Novel	Social	Novel	
NON	Mean	114.42	94.84	-	-	19.59	114.91	95.51	-	-	19.40
	SEM	5.30	8.12	-	-	11.78	7.96	7.70	-	-	15.22
AGG	Mean	121.48	97.66	-	-	23.82	106.41	95.02	-	-	11.38
	SEM	6.40	8.34	-	-	12.70	10.22	10.81	-	-	18.66
P value		0.41	0.81	-	-	0.81	0.52	0.97	-	-	0.74
n		7-8	7-8	-	-	7-8	8	8	-	-	8

<u>Test phase</u>		Duration (s)		CD-1 Target Normalized social score		Subtracted social score	Duration (s)		C57BL6/J Target Normalized social score		Subtracted social score
		Social	Novel	Social	Novel		Social	Novel	Social	Novel	
NON	Mean	166.49	93.31	1.50	1.08	73.17	153.47	95.64	1.41	1.04	57.83
	SEM	12.49	10.70	0.16	0.19	22.88	6.41	4.87	0.16	0.09	9.14
AGG	Mean	154.89	100.56	1.30	1.08	54.33	148.77	98.80	1.50	1.13	49.97
	SEM	5.03	7.04	0.09	0.12	10.61	7.60	8.42	0.16	0.14	14.21
P value		0.41	0.59	0.32	0.98	0.49	0.55	0.48	0.85	0.32	0.83
n		7-8	7-8	7-8	7-8	7-8	8	8	8	8	8

The behavioural data are shown as mean $\pm$ s.e.m. and analysed by unpaired Student's *t*-test. Significance at **P* < 0.05. The subtracted social score was derived by subtracting time in the social-paired chamber from the novel object-paired chamber during the test phase. Normalized social score is the ratio of time spent in the chamber of interest (social target or novel object) during the test phase over the pre-test phase.