nature protocols

PROTOCOL

https://doi.org/10.1038/s41596-020-0382-9

Check for updates

A standardized social preference protocol for measuring social deficits in mouse models of autism

Benjamin Rein, Kaijie Ma and Zhen Yan ✉

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by social communication deficits and other behavioral abnormalities. The three-chamber social preference test is often used to assess social deficits in mouse models of ASD. However, varying and often contradicting phenotypic descriptions of ASD mouse models can be found in the scientific literature, and the substantial variability in the methods used by researchers to assess social deficits in mice could be a contributing factor. Here we describe a standardized three-chamber social preference protocol, which is sensitive and reliable at detecting social preference deficits in several mouse models of ASD. This protocol comprises three phases that can all be completed within 1 d. The test mouse is first habituated to the apparatus containing two empty cups in the side chambers, followed by the pre-test phase in which the mouse can interact with two identical inanimate objects placed in the cups. During the test phase, the mouse is allowed to interact with a social stimulus (an unfamiliar wild-type (WT) mouse) contained in one cup and a novel non-social stimulus contained in the other cup. The protocol is thus designed to assess preference between social and non-social stimuli under conditions of equal salience. The broad implementation of the three-chamber social preference protocol presented here should improve the accuracy and consistency of assessments for social preference deficits associated with ASD and other psychiatric disorders.

Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by reduced or impaired social interaction, repetitive behaviors and/or restricted interests. ASD has been linked to a broad range of etiologies; >50% of ASD cases are thought to be caused by genetic variation[1], and genetic screenings have led to the identification and implication of numerous high-risk genes in ASD pathogenesis[2]. Transgenic mouse models carrying mutations in high-risk ASD genes or genetic loci, such as *Shank3*-deficiencies[3,4] or 16p11.2 copy number variations (CNVs; that is, deletion/duplication)[5,6], display ASD-related behavioral phenotypes[7–11], and represent powerful tools for elucidating neurobiological mechanisms that drive ASD pathogenesis.

Behavioral assays sensitive to social deficits are necessary for phenotypic verification of ASD models and for evaluation of therapeutic intervention strategies. The three-chamber social preference test, which assesses the animal's preference for a social stimulus over a non-social stimulus, is one of the most commonly used methods for evaluating sociability in mouse models of ASD[12]. However, numerous modifications have been made to this assay since it was initially described, resulting in an array of separate and distinct protocols across the ASD literature with dispersed usage[8,13–23]. The variety of testing methods has contributed to discrepancies between studies in the phenotypic descriptions of several ASD mouse models, with varying conclusions depending on the protocol used. For example, opposing phenotypes (that is, the presence or absence of social preference deficits) have been reported in *Shank2*$^{-/-}$ mice[16,17], *Shank3*$^{e4-9}$ mice[20,24], and *Shank3*$^{\Delta C/\Delta C}$ mice[7,18,19], with differing protocols used across studies. These findings suggest that the different three-chamber social preference test methodologies used may have differing sensitivities for detection of social deficits. To encourage consistent and reliable phenotyping of ASD-related social deficits in mice, we describe here a three-chamber social preference test protocol that offers robust detection of social preference deficits, and demonstrates enhanced sensitivity relative to a commonly used alternative approach.

Department of Physiology and Biophysics, Jacobs School of Medicine and Biomedical Sciences, State University of New York at Buffalo, Buffalo, NY, USA. ✉e-mail: zhenyan@buffalo.edu

Comparison between the social preference test variants

The three-chamber social preference assay we describe here evaluates the test mouse's preference for interacting with a social (S) stimulus versus a non-social (NS) stimulus (termed the 'three-phase S–NS' protocol). First, the test mouse is habituated to a three-chamber apparatus containing two empty cups, to reduce the salience of these objects (habituation phase). Next, two identical objects (paper balls) are placed within the cups, to familiarize the animal with the presence of objects contained within the cups (pre-test phase). Finally, a social stimulus (an age- and sex-matched WT mouse) is introduced under one cup and a novel non-social stimulus (wooden block) is placed under the other cup (test phase). The amount of time spent interacting with either stimulus is recorded in order to assess the animal's level of preference for the social stimulus over the non-social stimulus. This three-phase S–NS method is designed with the intention of minimizing variability caused by novelty-driven interactions with the cup and reliably isolate the animal's interest level towards a social stimulus versus a non-social stimulus. This protocol and similar protocols have been effective in identifying social preference deficits in several ASD mouse models, including *Shank3*-deficient mice, such as *Shank3*$^{+/\Delta C}$ mice[7,8], *Shank3*$^{\Delta C/\Delta C}$ mice[19], *Shank3*$^{e4-9}$ mice[20], *Shank2* knockout mice[16], forebrain *Cul3*-deficient mice (Cul3$^{f/-}$)[25] and 16p11.2 duplication mice (16p11.2$^{dp/+}$)[11].

The three-phase S–NS protocol represents a modification of a widely used three-chamber social testing method[12,26]. In this alternative variant, the test mouse is first habituated to an empty three-chamber apparatus (habituation phase). In the subsequent test phase, two cups are placed on opposite sides of the apparatus, one containing an age- and sex-matched WT mouse and the other being empty. Since this protocol compares the mouse's interaction with a social stimulus (S) versus an empty cup (E), it is referred as the 'two-phase S–E' method. Numerous variations of this method also exist. One commonly used method[24,27–30] is identical to the two-phase S–E method, but includes an additional habituation phase to only the center chamber before habituation to the entire apparatus. In another protocol[13,15,22], the test mouse is only habituated to the center chamber before the test phase. Another method habituates the test mouse to the apparatus containing two empty cups, and in the subsequent test phase, one cup contains a social stimulus, while the other remains empty[14,17,18].

Several papers have indicated that the three-phase S–NS protocol is more sensitive at detecting social deficits in ASD models than the two-phase S–E approach. Schmeisser et al.[17] reported normal social interaction time and social preference in *Shank2*$^{-/-}$ mice tested with the S–E approach. In contrast, Won et al.[16] used an S–NS three-chamber social preference protocol, and found that *Shank2*$^{-/-}$ mice spent significantly less time than WT animals interacting with the social stimulus. Additionally, homozygous mice with the deletion of *Shank3* exon 4–9 (*Shank3*$^{e4-9}$) were reported to display significant deficits in social preference when tested with the S–NS protocol[20]; however, a separate study reported normal social preference in *Shank3*$^{e4-9}$ mice when tested with an S–E protocol[24]. Furthermore, one study using an S–E approach reported the lack of three-chamber social preference deficits in *Shank3*$^{\Delta C/\Delta C}$ mice[18], whereas multiple studies using the three-phase S–NS approach did find robust social deficits in *Shank3*$^{\Delta C/\Delta C}$ mice[7,19] and male *Shank3*$^{+/\Delta C}$ mice[7,8,31].

The finding that disruption of *Shank3* is linked to autism in humans[3,4,32,33] and leads to social impairments in macaques[34,35] are consistent with the social deficits phenotypes in *Shank3*-deficient mice[7,8,15,19,20,31,36], but are in disagreement with the normal sociability in several lines of *Shank3* mutant mice detected with the S–E approach[18,24,37]. These results indicate that three-chamber social preference protocols utilizing a novel object placed under a cup, rather than an empty cup alone, are more sensitive to social preference deficits in ASD models. The three-phase S–NS protocol can also be used to examine social abnormalities relevant to the negative symptoms of schizophrenia (SZ), as shown previously with a three-armed platform containing empty wire cages on two arms[38]. The design of this method coincides with the protocol described here, in which the wire cup functions only as a component of the testing apparatus, while an inanimate object (a Lego mouse) was used as the non-social cue during the test phase. The results of ourselves and others suggest the three-phase S–NS three-chamber social preference protocol offers greatly improved sensitivity and robustness in revealing ASD-related social deficits.

Using the three-chamber S–NS approach, we did not observe social preference deficits in the mouse model of Phelan–McDermid Syndrome (PMS) with a complete deletion of *Shank3*, consistent with prior reports on its normal social interest[39] and social preference[40]. Although it is hard to explain this apparently distinct mouse phenotype from human and monkey studies, one possibility is the compensatory effects of other *Shank* family members in *Shank3*-deleted mice. While behavioral methods are critical in phenotypic characterization, the choice of mouse lines is also a key determining factor[41].

In this study, we found that the two-phase S–E protocol failed to detect social deficits in several mouse models of ASD. Nevertheless, we do not claim that this method is completely ineffective. Several studies following the S–E method or similar protocols have identified social preference deficits in ASD models, such as *Shank3B*$^{-/-}$ mice[15], *Pten* conditional knockout mice[42], and mice with homozygous deletion of the ASD-associated genes *Neuroligin-4* (ref. [13]) and *Cntnap2* (ref. [27]). However, another study using the same S–E protocol reported a lack of social deficits in *Cntnap2* knockout mice[30], suggesting that the S–E protocol may be prone to substantial variability. Thus, despite the capability of the S–E method to detect social preference deficits, it may be less sensitive to social deficits in mouse models that recapitulate ASD humans with haploinsufficiency of risk genes.

The negative results seen with the two-phase S–E method could be due to inherent design problems. The empty cup presented as the non-social stimulus also serves as a component of the social stimulus, as an identical cup is used to house the WT mouse. This may result in an inherent bias in favour of the social stimulus that contains both a novel social stimulus (mouse) and a novel non-social stimulus (cup), and is thus more salient than the non-social stimulus containing a cup alone. This inherent bias for the social stimulus driven by the design of the two-phase S–E protocol may mask the presence of social deficits in ASD models tested with this method. In addition, due to the lack of habituation to the empty cup, the test mouse is prone to engage in extended investigation of the cup, which may affect interaction time with either the social or non-social stimulus, promoting unplanned and unpreventable variability in sociability tests. We therefore encourage the use of the three-phase S–NS protocol, in order to improve the sensitivity, robustness and consistency of phenotypic screenings in mouse models of ASD.

Applications

To date, the three-chamber social preference protocol presented here has been used predominantly for phenotyping of social deficits in transgenic mouse models of ASD[7,11,16,19,20,25] and evaluating the effectiveness of treatment strategies[8,31,36]. However, this protocol may be appropriately applied in other contexts, including environmentally induced models of ASD[43,44], animals affected by physical or emotional stress[45] and functional studies of neurocircuitry controlling sociability[46].

Limitations

The described three-chamber sociability test (three-phase S–NS) offers robust sensitivity to the measurement of social preference; however, not all socially affected animals are guaranteed to exhibit deficits. Mice carrying deletion of 16p11.2 fail to display three-chamber social preference deficits[30], despite impairments in several other measurements of sociability, including social approach[10,47], male–female reciprocal social interactions[48] and ultrasonic vocalizations[49]. Therefore, this approach appears to be sensitive to context-specific deficits in preference for a social over a non-social stimulus, and should not be considered as a definitive indicator of the overall presence or absence of social deficits.

Sensory abnormalities are present in a large portion of children with ASD[50], and several mouse models of ASD exhibit various sensory phenotypes[51]. It is possible that sensory deficits may affect performances in social behavioral assays. However, 16p11.2 deletion mice, which are deaf and have reduced ultrasonic vocalizations[49], display normal social preference in three-chamber sociability tests[30], consistent with our findings here. Therefore, the presence of sensory deficits is not guaranteed to affect the social preference test. Nevertheless, examining visual, auditory and olfactory integrity is recommended.

Relative to two-phase S–E protocols, the method described here (three-phase S–NS) involves more rodent handling. However, the current protocol includes handling such as taking the test mouse out of the apparatus while cleaning and replacing objects between trials, which is potentially less disruptive to the animal's behavior than moving objects while the mouse is in the apparatus. Nevertheless, all animals should be handled gently to minimize stress.

Experimental design

This protocol is suitable for assessing social preference in all strains of mice. However, controls should be WT animals of the same strain, as baseline sociability may differ between mouse strains. Locomotion differences or motor deficits could be a confounding factor impacting test results. Animals of all ages may be tested, but controls must be age-matched, as sociability declines as animals

get older[52,53]. In addition, both male and female mice can be tested with this protocol, and controls should be sex-matched. We have reliably used this protocol on juvenile to adult animals (5–6 weeks old to 4–5 months old). For all experiments, WT littermates should be used as control groups. Unfamiliar age-, strain- and sex-matched WT mice should be used as the social stimulus. The use of genetically altered or otherwise socially impaired mice as the social stimulus may affect the sociability of the test mouse. Generally, two separate groups of mice should be used as the test mice and the stimulus mice. However, if mouse availability is limited, test mice (only WT) may be used as the social stimulus after they have completed their testing. All animals should be group housed before testing of sociability because single housing of animals will induce isolation stress and affect sociability. If highly variable social behavior is observed among animals within a single genotype, experiments should include more than 10 animals in the group, from at least three cohorts, in order to draw accurate conclusions regarding social preference differences between groups. Group sizes should be properly determined to avoid using too few or too many mice—see ARRIVE guidelines for details[54].

A typical experiment comprises a habituation trial followed by two testing trials. For some studies it is appropriate to repeat measures on the same animals, for example, following drug treatment. If undertaking such studies, there should be a break of at least 3 d before repeating assays and it is not necessary to repeat the habituation phase.

The protocol described here has been designed and optimized for use in mice. However, the three-chamber social preference test has also been performed in rats[55]. We thus think this protocol could be adapted for rats, with the use of an appropriately sized larger apparatus.

Materials

Biological materials

- Test mice. This protocol may be used for testing commercially obtained mice (for example, purchased from the Jackson Laboratory) or mouse lines newly created by research laboratories **! CAUTION** Mice used as controls must be of the same strain as the test group. All animals tested and compared must be similar in age, as interaction time with the social stimulus typically decreases as animals age. This protocol has been used reliably in juvenile to adult animals (5–6 weeks old to 4–5 months old). Mice are maintained on a 12 h light (6:00–18:00)/dark (18:00–6:00) cycle. They should be group housed with gender-matched conspecifics (2–4 mice per cage) and provided with standard enrichment. All experiments must receive approval from the relevant institutional review board and be conducted in accordance with local and national regulations. We obtained permission from State University of New York at Buffalo Institutional Animal Care and Use Committee (IACUC) to undertake the studies shown here **▲ CRITICAL** Animals that display major deficits in locomotion should not be used for this assay.
- Social stimulus mice. The mice used as the social stimulus must be age-, sex-, and strain-matched to test mice **▲ CRITICAL** Using social stimulus mice that differ from test mice on any of these parameters could impact test results **▲ CRITICAL** Do not use socially impaired mice (such as transgenic ASD models) as the social stimulus—this may reduce the amount of time the test mouse spends interacting with the social stimulus. The stimulus mouse should be unfamiliar to the test mice; do not use cagemates of the test mice.

Reagents

- 75% (vol/vol) ethanol (Decon Laboratories, cat. no. DSP-MD.43) diluted in ddH_2O **! CAUTION** Ethanol can carry some odors and may have a fixative action on the molecules that are present in urines and feces. A more thorough solution is to soak and wash with soap, rinse and dry **▲ CRITICAL** We use ethanol for cleaning the testing apparatus and objects between tests because it evaporates quickly and effectively removes odor.

Equipment

- Three-chambered apparatus. The apparatus we use has the following specifications: 102 cm (length) × 47 cm (width) × 45 cm (height). The walls of the apparatus are made of transparent Plexiglas. The two side chambers that the stimuli are placed in measure 33 cm (length) × 47 cm (width) **▲ CRITICAL** The three-chamber apparatus with side chambers should be large enough to permit the test mouse to explore the area outside of the cup. The use of a small three-chamber apparatus with limited exploration space in each side chamber may affect the measurement of the test animal's social

behaviors. Our apparatus is larger than many commercially available three-chamber apparatuses (for example, Ugo basile; San Diego Instruments), which have side chambers of ~20 cm (length) × 40 cm (width). This shorter chamber permits less space for exploration in the outside area surrounding the cup, which may interfere with accurate measurement of the time spent interacting with the social stimulus contained inside the cup.

- Cup or capsule to house the social or non-social stimulus. We use a wire pencil cup (chrome color, made of sturdy steel), 10.2 cm (diameter) × 10.8 cm (height), with ~1 cm gaps between bars, sufficient for animal interaction and sniffing (Spectrum Diversified Galaxy Pencil Holder; Spectrumdiversified.com) ▲ **CRITICAL** We recommend users keep an extra set of cups, so that one set may be cleaned while the other is in use.
- Glass bottle or other object to be placed on top of the cup to prevent the test mouse from climbing. We use a 250-ml glass bottle (Pyrex Reusable Media Storage Bottles; Fishersci.com)
- Inanimate object to be placed within the cup as the non-social stimulus. We use a square (2.5 cm) wooden block. Other objects, such as Lego structures of simple shapes, can also be used
- Two identical inanimate objects to use in the pre-test phase. Paper balls, which are simple to prepare and readily available, can be used ▲ **CRITICAL** The paper towel is crumpled by hand while wearing clean gloves to avoid transferring animal odor to the paper ball.
- Digital camcorder to record for subsequent scoring or rescoring of the test animal's behavior
- Video tracking and analysis softwares. We use Anymaze (Stoelting). Other animal-tracking software, such as EthoVision XT by Noldus or idTracker, can also be used

Procedure

Habituation ● Timing 10 min

1 Bring the test mice to the behavioral room and allow them to habituate for at least 1 h, with the room set to the testing conditions.
▲ **CRITICAL** Overhead lighting should be minimized to avoid anxiogenic effects that may affect social interaction time. Brightness should be measured in the center of all three chambers to ensure that the apparatus is evenly lit. Brightness should ideally be maintained at <50 lux.
? TROUBLESHOOTING

2 Place two clean, empty, inverted pencil cups into the three-chamber apparatus, each center approximately halfway between the midline and the far wall.
▲ **CRITICAL** The testing apparatus and cups should be cleaned and free of debris prior to starting any new test.
▲ **CRITICAL** Place a clean, empty 250 ml glass bottle upright on top of each cup to prevent the test mouse from climbing the cup. The bottle placed on top of each cup should be identical in size, shape and color.

3 Gently place the test mouse into the center of the apparatus. Start a timer and allow 10 min for the animal to explore freely while habituating to the apparatus and empty cups.
! CAUTION When transferring the test mouse from its home cage to the testing apparatus, the animal should be handled gently, preferably carried on one arm or the home cage lid. Do not suspend the animal by its tail while carrying it. Tests preceded by rough handling may be affected by animal stress.

4 Remove the test mouse from the apparatus and gently return to its home cage for a 5-min break.

5 Wipe down the apparatus, cups and bottles with 75% (vol/vol) ethanol to remove any residual odors that may affect subsequent tests.
■ **PAUSE POINT** At this stage, the animals can be returned to their home cages and the remaining trials may be optionally carried out on the following day. If this is done, on the next day, repeat step 1 before proceeding with the following procedure.

Pre-test ● Timing 10 min

6 Prepare two clean paper balls and place one under each inverted pencil cup. The two paper balls used should be of the same variety, as they are intended to represent identical objects. The paper balls should be placed in the center of the cup. When placing the cups into the chambers, leave enough space between the cup and the outer wall of the apparatus for the test mouse to explore the full periphery of the cup.
! CAUTION Wear clean gloves when crumpling and placing the paper balls under the pencil cups. Transferring odors onto the paper balls may affect the pre-test trial.

7 Gently place the test mouse into the apparatus. Start a timer and allow 10 min for the animal to familiarize with the presence of the objects contained within the cups.
8 Remove the test mouse from the apparatus and gently return it to its home cage for a 5-min break.
9 Remove the paper balls and wipe down the apparatus, cups and bottles with 75% (vol/vol) ethanol to remove any residual odors that might affect subsequent tests.

Social preference test ● Timing 10 min

10 Place an age-, sex- and strain-matched unfamiliar WT mouse under one cup, to serve as the social stimulus. Rough handling of the stimulus mouse may negatively affect social interactions with the test mouse; handle gently when placing the stimulus mouse into the cup.
▲ **CRITICAL** The stimulus mouse must be unfamiliar to the test mouse; do not use cagemates. The mouse used as the social stimulus should be interchanged regularly when conducting multiple tests to avoid exhaustion or social fatigue of the stimulus mouse.
11 Place a wooden block or another unfamiliar, inanimate object under the other cup to serve as the non-social stimulus.
▲ **CRITICAL** The location of the social or non-social stimulus in either side chamber should be counterbalanced between tests.
12 Place the test mouse into the apparatus containing the social and non-social stimuli. Start a timer and allow the mouse to explore for 10 min. The amount of time spent interacting with the social stimulus and the non-social stimulus should be recorded. This can be done manually by an experimentally blind researcher, or automatically by video tracking software such as Anymaze.
13 Return the test mouse and stimulus mouse to their respective home cages.
14 Remove the object and wipe down the apparatus, cups and bottles with 75% ethanol to remove any residual odors that may affect subsequent tests.

(Optional) Social novelty test ● Timing 10 min

15 Replace the non-social object from the previous trial with an unfamiliar WT mouse (age-, sex- and strain-matched) as the 'novel' social stimulus.
16 Place the test mouse into the apparatus containing the novel and familiar social stimuli. Start a timer and allow the animal to explore for 10 min. Record the amount of time spent interacting with each stimulus either manually or digitally.
17 Return the test mouse and both stimulus mice to their respective home cages.
18 Wipe down the apparatus, cups and media bottles with 75% ethanol to remove any residual odors which may affect subsequent tests.

Troubleshooting

Lighting

The lighting of the test room may affect sociability. Social interaction time is typically reduced when animals are tested in brighter conditions. If overhead lighting cannot be dimmed, a standing lamp may be used to light the room. However, the lamp must be kept at a safe distance from the testing chamber so as not to induce anxiogenic effects. The lighting must be consistent across all areas of the three-chamber apparatus to prevent animal preference for darker locations or chambers. In unevenly lit testing conditions, the animal will prefer dimmer areas, which could affect testing results.

Animal testing and scoring

Sometimes the test animal climbs the cup and remains at the top of the cup without interacting with the stimulus. Additionally, the software may fail to constantly track the test animal. In such cases, manual counting is more accurate (see 'Scoring methods' for details).

Expected values

When tested with the three-phase S–NS protocol, the average social interaction time for WT mice (of either sex, 6–8 weeks old) typically falls between ~125 and ~150 s for a 10-min testing session, though this may vary between 100 and 200 s depending on the strain and age of animals tested. The average non-social interaction time is typically between ~25 and ~50 s. The average social preference index for WT mice (C57BL6 background) should be 0.4–0.8. However, mice commonly exhibit natural variability in behavioral tendencies, even within a single strain or genotype, so values may fall within

a broader range than this. Due to this expected spectrum of social behaviours, it is emphasized that comparisons must be made between group averages, which include data from a sufficiently large number of mice of either genotype, and from several litters. If any animal presents a value that is determined to be a statistically significant outlier, this animal may be removed from the analysis.

Housing effects

All test mice should be group housed, as single housing can produce severe detrimental effects on sociability and other behaviors due to social isolation stress[56]. Furthermore, housing mice with conspecifics of different genotypes can affect social behavior. Mouse models of ASD may be more likely to assume submissive roles in social hierarchies, as demonstrated in neuroligin-3 deficient (*Nlgn3*$^{y/-}$) mice[57], which may produce defeat-related social deficits. Indeed, male *Nlgn3*$^{y/-}$ mice housed with WT animals display more severe social deficits than those housed with genotype-matched conspecifics[57]. Interestingly, raising WT mice with *Nlgn3*$^{y/-}$ mice also compromises the sociability of WT. The negative impact of mixed-genotype housing on social behavior has been similarly reported in 16p11.2$^{+/-}$ mice[58]. There is also a report showing that enhancing environmental enrichment within animal housing improves sociability in valproic acid-exposed autism model mice[59]. These findings highlight the importance of carefully controlling housing conditions in order to produce accurate measurements of social behavior in ASD models.

Scoring methods

Scoring can be undertaken manually or using automated behavior-tracking software. The key information is the duration of direct interactions of the test mouse with the social or non-social stimulus. We usually use automated scoring of the three-chamber social preference test with Anymaze behavior-tracking software (Stoelting). The area directly surrounding the cup is designated as a zone of interest, and the amount of time spent in the zone by the test mouse is measured. This method of scoring therefore measures the amount of time the test animal spends in close proximity to the cup (distance of animal head to cup edge: ≤3.5 cm), rather than specifically measuring time spent sniffing or engaging with the social stimulus. Automated scoring may produce inaccurate conclusions, for example, if animals remain in the vicinity of the cup without interacting with the social or non-social stimulus. For this reason, manual scoring may be required to verify the scores. Automated scoring is also susceptible to software errors if the animal is not properly tracked. All videos should therefore be reviewed to verify that the animal has been tracked well.

If manual scoring is performed, all scoring should be performed by a researcher blinded to animal genotype and/or treatment. For manual scoring purposes, behaviors that are typically counted as interactions include: directly interacting with the stimulus mouse or non-social object between the wire bars of the pencil cup; sniffing the base of the cup containing the stimulus; interacting with parts of the stimulus that are protruding from the cup, such as the tail of the stimulus mouse; and actively attending to (sniffing/facing) the stimulus while climbing the cup. Behaviors that are not counted include: interacting with the bottle on top of the cup; standing near the cup without attending to (sniffing/facing) the cup or the contained stimulus; and self-grooming in the proximity of the cup. While it is helpful to use clearly defined scoring parameters, experimenters may differ in their assessment of behavior and therefore produce different values. Thus, all videos generated within a single experiment should be scored by the same experimenter to minimize human error. We recommend automated scoring followed by manual correction, which gives the most accurate results. Supplementary Videos 1 and 2 show examples of a WT and a *Shank3*-deficient mouse in the social preference test phase, with added commentary. Additionally, Supplementary Table 1 provides a list of various observable behaviors throughout these two videos, and indicates how they should be manually scored.

Statistical analysis

All behavioral testing should be performed on at least three independent cohorts. Interaction time with the social stimulus (T_S) and non-social stimulus (T_{NS}) is quantified. For comparisons between WT versus mutant, a two-way ANOVA should be performed with comparisons between all four values (T_S in WT, T_{NS} in WT, T_S in mutant, T_{NS} in mutant), followed by post hoc Bonferroni tests for multiple comparisons within and between groups. In addition, social preference indexes, $I_{SP} = (T_S - T_{NS})/(T_S + T_{NS})$, are compared between groups using two-tailed Student's *t*-tests. All datasets should be tested for normality using Shapiro–Wilk tests, and data that fail normality tests are compared with non-parametric tests, such as the Mann–Whitney *U*-test.

For a genotype where T_S is significantly greater than T_{NS}, this represents the existence of social preference. A genotype showing a significant reduction of both T_S and I_{SP} relative to the WT group warrants the interpretation that social deficits, including the impairment in social engagement, social interest, social interaction and social preference, are manifested. If T_S is unchanged, and only I_{SP} is significantly reduced (due to the increased T_{NS}) in the mutant group, this suggests the presence of relatively mild social abnormality, reflected by the impairment in social preference.

Repeated measures

The three-phase S–NS three-chamber social preference protocol can be performed repeatedly on the same mice with consistent results. Several of our studies have included repeated testing at progressive intervals in control and treatment groups across time points, to test the longitudinal therapeutic efficacy[8,31,36]. The three-phase S–NS protocol is well suited for repeated testing, as a new object may be placed inside the cup in the test phase during repeated measurements, thereby preserving the novelty of the non-social stimulus. This represents an additional advantage of the three-phase S-NS protocol over the two-phase S–E protocol, as the empty cup becomes familiar after a single test and cannot be considered a novel non-social stimulus in subsequent testing unless different types of cups are used each time.

When performing repeated testing with the three-phase S–NS method, use a novel social stimulus in subsequent tests following the initial assessment. Additionally, in subsequent days following the initial testing, the habituation phase to the empty cups may be omitted, and performing only two phases (pre-test and social preference) is sufficient.

Typical results seen using the three-phase S–NS social preference protocol

In this section, we discuss examples of results that have been obtained by following the three-phase S–NS protocol (Fig. 1a), and demonstrate its sensitivity in detecting ASD-related social preference deficits in several distinct transgenic mouse models of ASD. Detailed statistical information for all data are included in the Source Data for Figs. 1 and 2.

Shank3, which encodes a postsynaptic scaffolding protein located at glutamatergic synapses, is among the strongest genetic risk factors for ASD[3,4] and plays a causal role in PMS[60]. Exon 21, the largest coding region of *SHANK3*, has the most variants and mutations in humans with ASD[3,4,32,33]. We tested heterozygous mice carrying exon 21-deleted *Shank3* gene, which results in the truncated form of Shank3 protein lacking the C-terminal region (Shank3$^{+/\Delta C}$), mimicking the human ASD-linked disruption of *SHANK3* exon 21 (ref. [32]). The 6- to 8-week-old male Shank3$^{+/\Delta C}$ mice spent significantly less time than WT littermates investigating the social stimulus, and did not exhibit a significant preference for the social stimulus over the non-social stimulus (Fig. 1b, WT: $n = 8$; Shank3$^{+/\Delta C}$: $n = 14$, $F_{1,40(\text{interaction})} = 10.0$, $P < 0.01$, two-way ANOVA). Shank3$^{+/\Delta C}$ mice correspondingly displayed a significantly lower social preference index than WT mice (Fig. 1c, $t_{(20)} = 3.94$, $P < 0.01$, unpaired *t*-test), indicating social interaction deficits. Two videos showing one male WT and one male Shank3$^{+/\Delta C}$ mouse in the social preference test phase with the three-phase S–NS method are included as Supplementary Videos 1 and 2.

We then tested 6- to 8-week-old female Shank3$^{+/\Delta C}$ mice with the three-phase S–NS method. Unlike male Shank3$^{+/\Delta C}$ mice, female Shank3$^{+/\Delta C}$ spent significantly more time interacting with the social stimulus than the non-social stimulus (Fig. 1e, $n = 9$ mice/group, $F_{1,32(\text{interaction})} = 0.4$, $P > 0.5$, two-way ANOVA), and exhibited a social preference index similar to female WT animals (Fig. 1f, $t_{(16)} = 1.1$, $P > 0.2$, unpaired *t*-test). This suggests that heterozygous *Shank3* exon 21 deletion confers sociability deficits that are restricted to male mice, and that the three-phase S–NS method is capable of isolating sex-specific deficits within a single genotype.

CNVs of the human 16p11.2 gene locus are among the strongest genetic risk factors for ASD[5,6,61]. Mice carrying deletion or duplication of the 16p11.2 murine orthologue exhibit behavioral features of neurodevelopmental disorders including ASD-related social deficits[10,11,47,49,62,63]. We tested male and female 6- to 8-week-old 16p11.2 duplication mice (16p11.2$^{dp/+}$) and WT littermates using the three-phase S–NS method, and found that 16p11.2$^{dp/+}$ mice spent significantly less time than WT animals interacting with the social stimulus, and failed to display a significant preference for the social stimulus over the non-social stimulus (Fig. 1h, WT: $n = 10$; 16p11.2$^{dp/+}$: $n = 12$, $F_{1,40(\text{interaction})} = 11.5$, $P < 0.01$, two-way ANOVA). Correspondingly, the social preference index for 16p11.2$^{dp/+}$ mice was significantly reduced relative to WT mice (Fig. 1i, $t_{(20)} = 2.5$, $P < 0.05$, unpaired *t*-test).

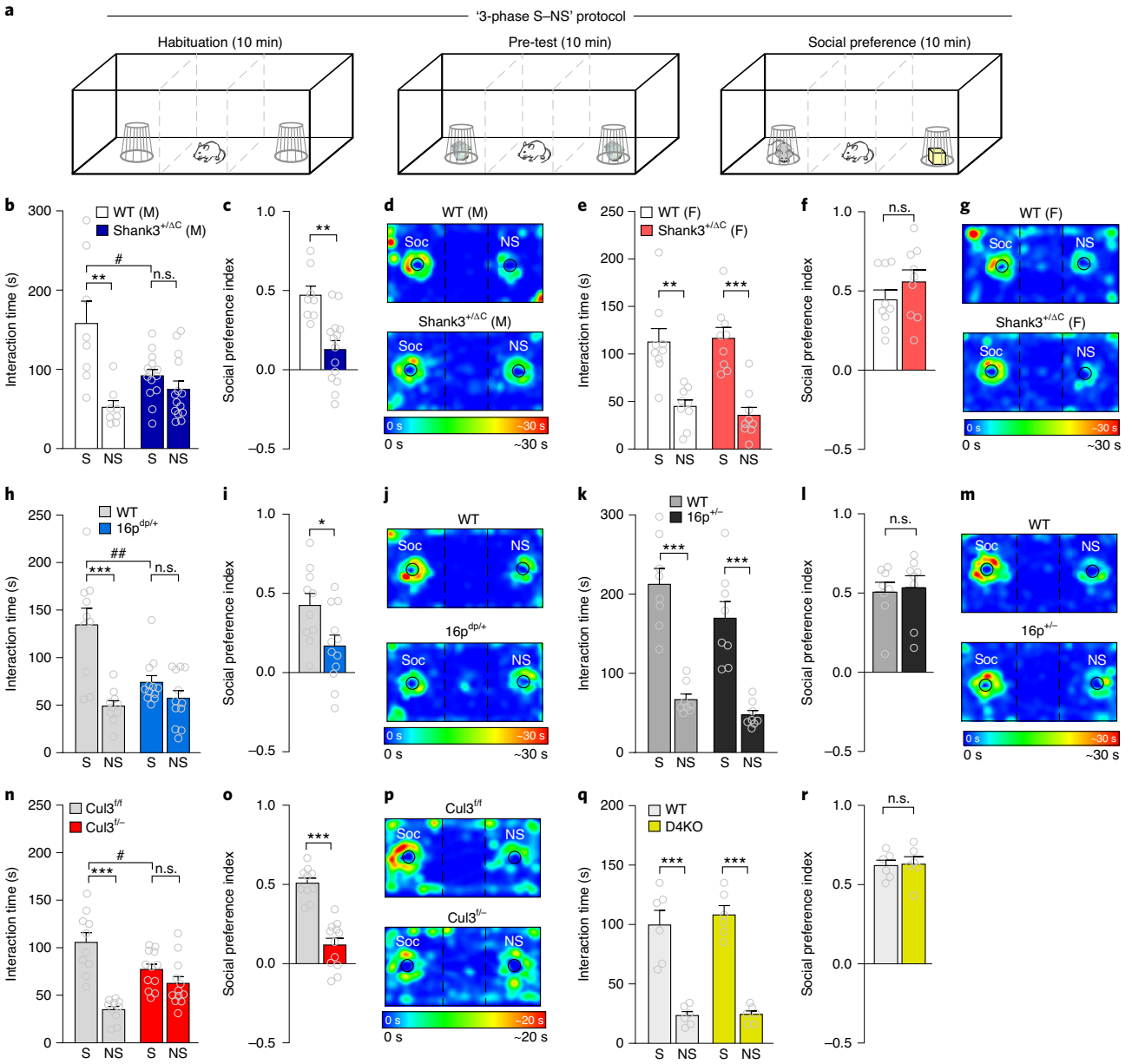


Fig. 1 | Social behavioral data obtained from several transgenic mouse models using the three-phase S–NS protocol. a, Graphic depicting the three-phase S–NS protocol, consisting of a 10-min habituation phase to the apparatus containing two empty cups, a 10-min pre-test phase in which two identical objects (paper balls) are placed under the cups, and a 10-min social preference test phase in which one cup contains a social (S) stimulus (age- and sex-matched WT mouse) and the other contains a non-social (NS) stimulus (wooden block). **b**,**e**,**h**,**k**,**n**,**q**, Bar graphs showing the amount of time spent interacting with the social stimulus (S) or non-social stimulus (NS) in male WT versus Shank3$^{+/\Delta C}$ mice (**b**), female WT versus Shank3$^{+/\Delta C}$ mice (**e**), WT versus 16p11.2$^{dp/+}$ mice (**h**), WT versus 16p11.2$^{+/-}$ mice (**k**), Cul3$^{f/f}$ versus Cul3$^{f/-}$ mice (**n**) and WT versus D4KO mice (**q**). Both sexes were used in **g**,**j**,**m**,**p**,**c**,**f**,**i**,**l**,**o**,**r**, Bar graphs comparing the social preference index of individual mouse lines. **d**,**g**,**j**,**m**,**p**, Representative heat maps illustrating the topographical time distribution in social preference tests of individual mouse lines. All data are presented as mean ± s.e.m. For all figures: n.s., not significant, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.0001$, S versus NS; $^{\#}P < 0.05$, $^{\#\#}P < 0.01$, WT versus mutant (social time). Note, the results in **b**–**d**, **h**–**j** and **n**–**p** are consistent with prior findings in refs. [7,8,11,25,31,36]. All animal studies were performed with the approval of the IACUC of the State University of New York at Buffalo.Source data

A previous characterization of 16p11.2 deletion mice (16p11.2$^{+/-}$) found that they display normal sociability when tested with the two-phase S–E method[30]. We thus tested 16p11.2$^{+/-}$ mice (males and females, 6–7 weeks old) with the three-phase S–NS method to determine whether they may exhibit social deficits with this more sensitive approach. Similar to WT mice, 16p11.2$^{+/-}$ mice spent significantly more time interacting with the social stimulus than the non-social stimulus (Fig. 1k, $n = 8$ mice/group, $F_{1,28(\text{interaction})} = 0.6$, $P = 0.4$, two-way ANOVA), and the social preference index was not significantly altered (Fig. 1l, $n = 8$ mice/group, $U = 24$, $P = 0.43$, Mann–Whitney U-test).

These findings confirm that 16p11.2$^{+/-}$ mice do not display three-chamber social preference deficits, despite exhibiting impairments in various other sociability assays[10,48,49,63].

We next tested 6- to 8-week-old male and female mice with forebrain-specific deletion of the high-risk ASD gene *Cul3* (Cul3$^{f/-}$)[25]. Unlike Cul3$^{f/f}$ controls, Cul3$^{f/-}$ mice failed to show a significant preference for the social over the non-social stimulus (Fig. 1n, Cul3$^{f/f}$: $n = 10$; Cul3$^{f/-}$: $n = 12$, $F_{1,40(\text{interaction})} = 16.2$, $P < 0.001$, two-way ANOVA), and exhibited a significantly reduced social preference index (Fig. 1o, $t_{(20)} = 7.2$, $P < 0.0001$, unpaired *t*-test), indicating the presence of social deficits.

The dopamine D4 receptor (D4R) is implicated in SZ[64], and D4 receptor knockout mice (D4KO) display hypersensitivity to psychomotor stimulants[65] and stress-induced SZ-related phenotypes[66]. However, sociability is unimpaired in these animals[66]. We thus utilized D4KO mice as a negative control to verify the reliability of the three-phase S–NS method in detecting social deficits without yielding false positives in socially unaffected transgenic models. Similar to WT mice, D4KO mice spent significantly more time interacting with the social stimulus than the non-social stimulus (Fig. 1q, $n = 6$ mice/group, $F_{1,20(\text{interaction})} = 0.2$, $P = 0.6$, two-way ANOVA), and did not differ from WT animals in their social preference index (Fig. 1r, $t_{(10)} = 0.2$, $P = 0.9$, unpaired *t*-test), confirming the lack of social deficits in D4KO mice.

Collectively, these results indicate that the 3-phase S–NS protocol has robust sensitivity in revealing social deficits in distinct mouse models of ASD (male Shank3$^{+/\Delta C}$, 16p11.2$^{dp/+}$ and Cul3$^{f/-}$). Moreover, this method retains high reliability in confirming the lack of social preference deficits in multiple mouse lines (female Shank3$^{+/\Delta C}$, 16p11.2$^{+/-}$ and D4KO).

As mentioned earlier, the three-phase S–NS protocol may optionally be augmented to assess preference for a novel social stimulus over a familiar social stimulus. We do not include example data from the social novelty preference phase here, and interested readers are encouraged to refer to our previous papers on social novelty preference data for *Shank3*$^{+/\Delta C}$ and 16p11.2$^{dp/+}$ ASD mouse models[7,11].

Comparison with the two-phase S–E social preference protocol

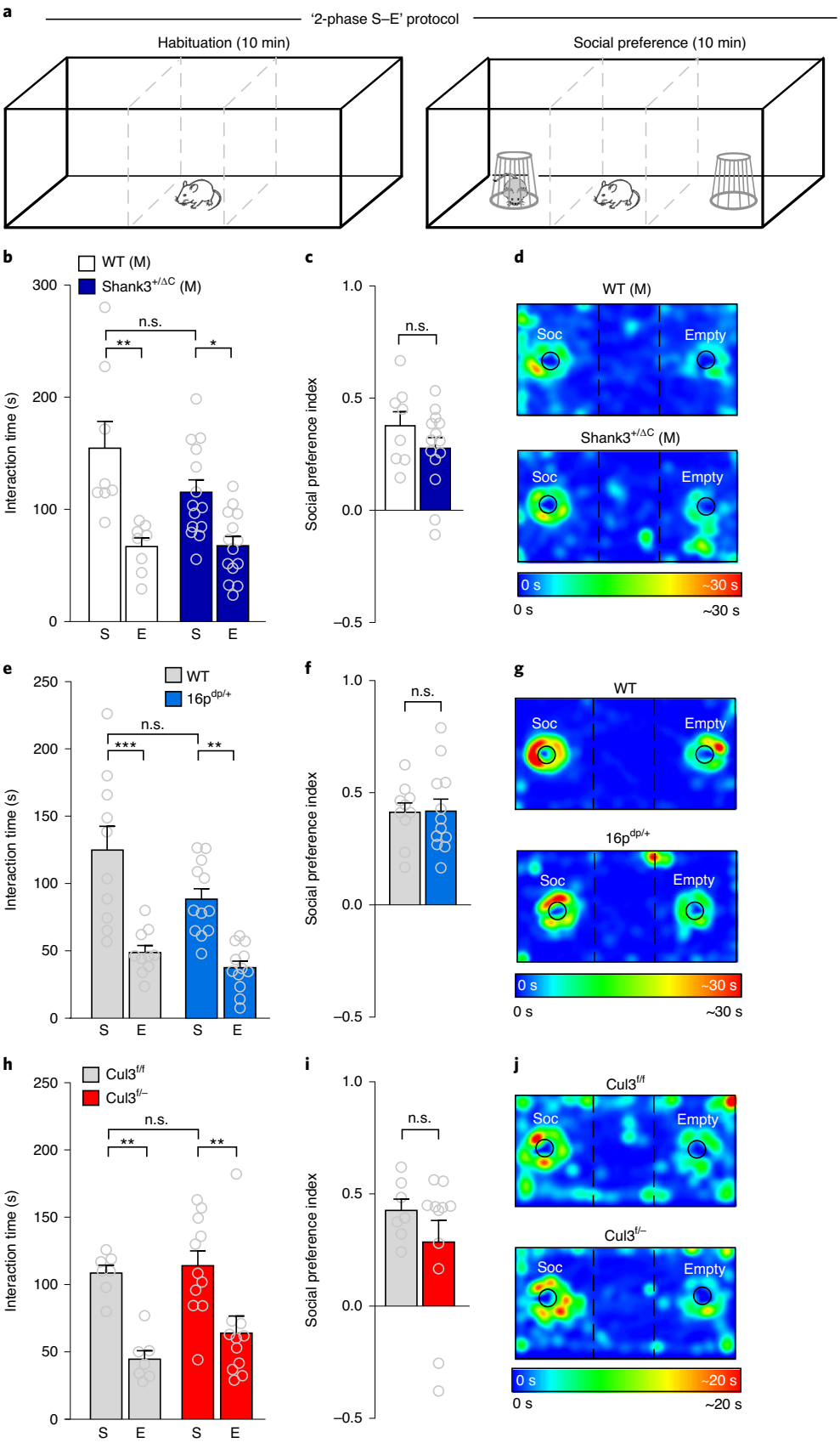
To compare differences in sensitivity between the three-phase S–NS protocol and the widely used two-phase S–E protocol, we also tested the same mouse models of ASD with the two-phase S–E protocol (Fig. 2a). Using this testing method, male Shank3$^{+/\Delta C}$ mice did not differ from WT animals in the amount of time spent interacting with the social stimulus, and showed a significant preference for the social stimulus over the empty cup (Fig. 2b, WT: $n = 8$; Shank3$^{+/\Delta C}$: $n = 14$, $F_{1,40(\text{interaction})} = 2.4$, $P = 0.13$, two-way ANOVA). Additionally, the social preference index did not differ between male WT and Shank3$^{+/\Delta C}$ mice (Fig. 2c, $t_{(20)} = 1.7$, $P = 0.10$, unpaired *t*-test). These findings indicate that the two-phase S–E protocol fails to reveal social preference deficits in male Shank3$^{+/\Delta C}$ mice, contrary to the findings from the three-phase S–NS protocol (Fig. 1b–d). Our results suggest that contradicting phenotypic descriptions of *Shank3*-deficient mice in the existing literature may be due to different testing methods.

Further testing of 16p11.2$^{dp/+}$ and Cul3$^{f/-}$ mice (males and females, 6–8 weeks old) with the two-phase S–E protocol indicated that they spent significantly more time interacting with the social stimulus than the empty cup (Fig. 2e, WT: $n = 10$, 16p11.2$^{dp/+}$: $n = 12$, $F_{1,40(\text{interaction})} = 1.7$, $P = 0.2$, two-way ANOVA; Fig. 2h, Cul3$^{f/f}$: $n = 7$, Cul3$^{f/-}$: $n = 11$, $F_{1,32(\text{interaction})} = 0.4$, $P = 0.5$, two-way ANOVA), and their social preference indexes did not differ from those of WT controls (Fig. 2f, $t_{(20)} = 0.07$, $P = 0.9$, unpaired *t*-test; Fig. 2i, $U = 33$, $P = 0.65$, Mann–Whitney *U*-test).

Collectively, these findings indicate that three distinct mouse models of ASD, which display clear social deficits using the three-phase S–NS protocol (male Shank3$^{+/\Delta C}$, 16p11.2$^{dp/+}$ and Cul3$^{f/-}$), fail to show social preference deficits using the two-phase S-E protocol, suggesting that the three-phase S–NS protocol offers higher sensitivity to detect social deficits in ASD models. We thus propose that the adoption of this method should be prioritized to maximize the accuracy of phenotypic behavioral screenings.

Timing

Animals need to be transferred to the behavioral room at least 60 min prior to testing. This protocol requires ~40 min to perform per animal if all three trials are undertaken (three 10-min testing trials, with two 5-min intervals between trials for animal resting and apparatus cleaning). The 10-min habituation trial may optionally be completed 1 d before the pre-test and social preference test.


a
'2-phase S–E' protocol
Habituation (10 min)
Social preference (10 min)
b
WT (M)
Shank3+/ΔC (M)
Interaction time (s)
n.s.
**
*
S E S E
c
Social preference index
n.s.
d
WT (M)
Soc
Empty
Shank3+/ΔC (M)
0 s
~30 s
e
WT
16pdp/+
Interaction time (s)
n.s.
***
**
f
Social preference index
n.s.
g
WT
16pdp/+
h
Cul3f/f
Cul3f/–
Interaction time (s)
n.s.
**
i
Social preference index
n.s.
j
Cul3f/f
Cul3f/–
0 s
~20 s

◀ **Fig. 2 | Social behavioral data obtained from several ASD mouse models using the two-phase S–E protocol.** **a**, Graphic depicting the two-phase S–E protocol, consisting of a 10-min habituation phase to the empty apparatus and a 10-min social preference test phase in which a social stimulus (age- and sex-matched WT mouse under cup) and non-social stimulus (empty cup) are introduced. **b**,**e**,**h**, Bar graphs showing the amount of time spent interacting with the social stimulus (S) or empty cup (E) in male WT versus Shank3$^{+/\Delta C}$ mice (**b**), WT versus 16p11.2$^{dp/+}$ mice (**e**) and Cul3$^{f/f}$ versus Cul3$^{f/-}$ mice (**h**). Both sexes were used in **e** and **h**. **c**,**f**,**i**, Bar graphs comparing the social preference index of individual mouse lines. **d**,**g**,**j**, Representative heat maps illustrating the topographical time distribution in social preference tests of individual mouse lines. All data are presented as mean ± s.e.m. For all figures: n.s., not significant, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.0001$. All animal studies were performed with the approval of the IACUC of the State University of New York at Buffalo.Source data

When performing repeated measures on the same animals, allow at least 3 d between assays. It takes ~25 min for each animal in subsequent days (habituation phase omitted, two 10-min testing trials, with one 5-min interval).

Reporting Summary

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

Data availability

An Excel file containing all the statistical data for the two figures is included in the Source Data for Figs. 1 and 2.

References

1. De Rubeis, S. & Buxbaum, J. D. Genetics and genomics of autism spectrum disorder: embracing complexity. *Hum. Mol. Genet.* **24**, R24–R31 (2015).
2. Satterstrom, F. K. et al. Large-scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism. *Cell* **180**, 568–584 e523 (2020).
3. Leblond, C. S. et al. Meta-analysis of SHANK mutations in autism spectrum disorders: a gradient of severity in cognitive impairments. *PLoS Genet.* **10**, e1004580 (2014).
4. Betancur, C. & Buxbaum, J. D. SHANK3 haploinsufficiency: a "common" but underdiagnosed highly penetrant monogenic cause of autism spectrum disorders. *Mol. Autism* **4**, 17 (2013).
5. Kumar, R. A. et al. Recurrent 16p11.2 microdeletions in autism. *Hum. Mol. Genet.* **17**, 628–638 (2008).
6. Weiss, L. A. et al. Association between microdeletion and microduplication at 16p11.2 and autism. *N. Engl. J. Med.* **358**, 667–675 (2008).
7. Duffney, L. J. et al. Autism-like deficits in shank3-deficient mice are rescued by targeting actin regulators. *Cell Rep.* **11**, 1400–1413 (2015).
8. Qin, L. et al. Social deficits in Shank3-deficient mouse models of autism are rescued by histone deacetylase (HDAC) inhibition. *Nat. Neurosci.* **21**, 564–575 (2018).
9. Rein, B., Yan, Z. & Wang, Z. J. Diminished social interaction incentive contributes to social deficits in mouse models of autism spectrum disorder. *Genes Brain Behav.* **19**, e12610 (2019).
10. Wang, W. et al. Chemogenetic activation of prefrontal cortex rescues synaptic and behavioral deficits in a mouse model of 16p11.2 deletion syndrome. *J. Neurosci.* **38**, 5939–5948 (2018).
11. Rein, B. et al. Reversal of synaptic and behavioral deficits in a 16p11.2 duplication mouse model via restoration of the GABA synapse regulator Npas4. *Mol. Psychiatry* https://doi.org/10.1038/s41380-020-0693-9 (2020).
12. 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).
13. Jamain, S. et al. Reduced social interaction and ultrasonic communication in a mouse model of monogenic heritable autism. *Proc. Natl Acad. Sci. USA* **105**, 1710–1715 (2008).
14. Page, D. T., Kuti, O. J., Prestia, C. & Sur, M. Haploinsufficiency for Pten and Serotonin transporter cooperatively influences brain size and social behavior. *Proc. Natl Acad. Sci. USA* **106**, 1989–1994 (2009).
15. Peca, J. et al. Shank3 mutant mice display autistic-like behaviours and striatal dysfunction. *Nature* **472**, 437–442 (2011).
16. Won, H. et al. Autistic-like social behaviour in Shank2-mutant mice improved by restoring NMDA receptor function. *Nature* **486**, 261–265 (2012).
17. Schmeisser, M. J. et al. Autistic-like behaviours and hyperactivity in mice lacking ProSAP1/Shank2. *Nature* **486**, 256–260 (2012).
18. Kouser, M. et al. Loss of predominant Shank3 isoforms results in hippocampus-dependent impairments in behavior and synaptic transmission. *J. Neurosci.* **33**, 18448–18468 (2013).
19. Bidinosti, M. et al. CLK2 inhibition ameliorates autistic features associated with SHANK3 deficiency. *Science* **351**, 1199–1203 (2016).

20. Wang, X. et al. Synaptic dysfunction and abnormal behaviors in mice lacking major isoforms of Shank3. *Hum. Mol. Genet.* **20**, 3093–3108 (2011).
21. Drapeau, E., Dorr, N. P., Elder, G. A. & Buxbaum, J. D. Absence of strong strain effects in behavioral analyses of Shank3-deficient mice. *Dis. Model. Mech.* **7**, 667–681 (2014).
22. Radyushkin, K. et al. Neuroligin-3-deficient mice: model of a monogenic heritable form of autism with an olfactory deficit. *Genes Brain Behav.* **8**, 416–425 (2009).
23. Jung, E. M. et al. Arid1b haploinsufficiency disrupts cortical interneuron development and mouse behavior. *Nat. Neurosci.* **20**, 1694–1707 (2017).
24. Yang, M. et al. Reduced excitatory neurotransmission and mild autism-relevant phenotypes in adolescent Shank3 null mutant mice. *J. Neurosci.* **32**, 6525–6541 (2012).
25. Rapanelli, M. et al. Behavioral, circuitry, and molecular aberrations by region-specific deficiency of the high-risk autism gene Cul3. *Mol. Psychiatry* https://doi.org/10.1038/s41380-019-0498-x (2019).
26. Nadler, J. J. et al. Automated apparatus for quantitation of social approach behaviors in mice. *Genes Brain Behav.* **3**, 303–314 (2004).
27. Penagarikano, O. et al. Absence of CNTNAP2 leads to epilepsy, neuronal migration abnormalities, and core autism-related deficits. *Cell* **147**, 235–246 (2011).
28. Grayton, H. M., Missler, M., Collier, D. A. & Fernandes, C. Altered social behaviours in neurexin 1alpha knockout mice resemble core symptoms in neurodevelopmental disorders. *PLoS ONE* **8**, e67114 (2013).
29. Silverman, J. L. et al. Sociability and motor functions in Shank1 mutant mice. *Brain Res.* **1380**, 120–137 (2011).
30. Brunner, D. et al. Comprehensive analysis of the 16p11.2 deletion and null Cntnap2 mouse models of autism spectrum disorder. *PLoS ONE* **10**, e0134572 (2015).
31. Wang, Z. J. et al. Amelioration of autism-like social deficits by targeting histone methyltransferases EHMT1/2 in Shank3-deficient mice. *Mol. Psychiatry* https://doi.org/10.1038/s41380-019-0351-2 (2019).
32. Durand, C. M. et al. Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. *Nat. Genet.* **39**, 25–27 (2007).
33. Moessner, R. et al. Contribution of SHANK3 mutations to autism spectrum disorder. *Am. J. Hum. Genet.* **81**, 1289–1297 (2007).
34. Tu, Z. et al. CRISPR/Cas9-mediated disruption of SHANK3 in monkey leads to drug-treatable autism-like symptoms. *Hum. Mol. Genet.* **28**, 561–571 (2019).
35. Zhou, Y. et al. Atypical behaviour and connectivity in SHANK3-mutant macaques. *Nature* **570**, 326–331 (2019).
36. Ma, K. et al. Histone deacetylase inhibitor MS-275 restores social and synaptic function in a Shank3-deficient mouse model of autism. *Neuropsychopharmacology* **43**, 1779–1788 (2018).
37. Speed, H. E. et al. Autism-associated insertion mutation (InsG) of Shank3 Exon 21 causes impaired synaptic transmission and behavioral deficits. *J. Neurosci.* **35**, 9648–9665 (2015).
38. Bitanihirwe, B. K., Peleg-Raibstein, D., Mouttet, F., Feldon, J. & Meyer, U. Late prenatal immune activation in mice leads to behavioral and neurochemical abnormalities relevant to the negative symptoms of schizophrenia. *Neuropsychopharmacology* **35**, 2462–2478 (2010).
39. Wang, X. et al. Altered mGluR5-Homer scaffolds and corticostriatal connectivity in a Shank3 complete knockout model of autism. *Nat. Commun.* **7**, 11459 (2016).
40. Drapeau, E., Riad, M., Kajiwara, Y. & Buxbaum, J. D. Behavioral phenotyping of an improved mouse model of Phelan-McDermid syndrome with a complete deletion of the Shank3 gene. *eNeuro* **5**, ENEURO.0046-18.2018 (2018).
41. Jiang, Y. H. & Ehlers, M. D. Modeling autism by SHANK gene mutations in mice. *Neuron* **78**, 8–27 (2013).
42. Kwon, C. H. et al. Pten regulates neuronal arborization and social interaction in mice. *Neuron* **50**, 377–388 (2006).
43. Eissa, N. et al. The histamine H3R antagonist DL77 attenuates autistic behaviors in a prenatal valproic acid-induced mouse model of autism. *Sci. Rep.* **8**, 13077 (2018).
44. Imai, K. et al. Administration of molecular hydrogen during pregnancy improves behavioral abnormalities of offspring in a maternal immune activation model. *Sci. Rep.* **8**, 9221 (2018).
45. Zain, M. A., Pandy, V., Majeed, A. B. A., Wong, W. F. & Mohamed, Z. Chronic restraint stress impairs sociability but not social recognition and spatial memoryin C57BL/6J mice. *Exp. Anim.* **68**, 113–124 (2019).
46. Carta, I., Chen, C. H., Schott, A. L., Dorizan, S. & Khodakhah, K. Cerebellar modulation of the reward circuitry and social behavior. *Science* **363**, eaav0581 (2019).
47. Arbogast, T. et al. reciprocal effects on neurocognitive and metabolic phenotypes in mouse models of 16p11.2 deletion and duplication syndromes. *PLoS Genet.* **12**, e1005709 (2016).
48. Stoppel, L. J. et al. R-Baclofen reverses cognitive deficits and improves social interactions in two lines of 16p11.2 deletion mice. *Neuropsychopharmacology* **43**, 513–524 (2018).
49. Yang, M. et al. 16p11.2 deletion syndrome mice display sensory and ultrasonic vocalization deficits during social interactions. *Autism Res.* **8**, 507–521 (2015).
50. Marco, E. J., Hinkley, L. B., Hill, S. S. & Nagarajan, S. S. Sensory processing in autism: a review of neurophysiologic findings. *Pediatr. Res.* **69**, 48R–54R (2011).
51. Balasco, L., Provenzano, G. & Bozzi, Y. Sensory abnormalities in autism spectrum disorders: a focus on the tactile domain, from genetic mouse models to the clinic. *Front. Psychiatry* **10**, 1016 (2019).

52. Salchner, P., Lubec, G. & Singewald, N. Decreased social interaction in aged rats may not reflect changes in anxiety-related behaviour. *Behav. Brain Res.* **151**, 1–8 (2004).
53. Boyer, F., Jaouen, F., Ibrahim, E. C. & Gascon, E. Deficits in social behavior precede cognitive decline in middle-aged mice. *Front. Behav. Neurosci.* **13**, 55 (2019).
54. Kilkenny, C., Browne, W. J., Cuthill, I. C., Emerson, M. & Altman, D. G. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS Biol.* **8**, e1000412 (2010).
55. Bronzuoli, M. R. et al. Neuroglia in the autistic brain: evidence from a preclinical model. *Mol. Autism* **9**, 66 (2018).
56. Kappel, S., Hawkins, P. & Mendl, M. T. To group or not to group? Good practice for housing male laboratory mice. *Animals (Basel)* **7**, 88 (2017).
57. Kalbassi, S., Bachmann, S. O., Cross, E., Roberton, V. H. & Baudouin, S. J. Male and female mice lacking neuroligin-3 modify the behavior of their wild-type littermates. *eNeuro* **4**, ENEURO.0145-17.2017 (2017).
58. Yang, M., Lewis, F., Foley, G. & Crawley, J. N. In tribute to Bob Blanchard: divergent behavioral phenotypes of 16p11.2 deletion mice reared in same-genotype versus mixed-genotype cages. *Physiol. Behav.* **146**, 16–27 (2015).
59. Yamaguchi, H. et al. Environmental enrichment attenuates behavioral abnormalities in valproic acid-exposed autism model mice. *Behav. Brain Res.* **333**, 67–73 (2017).
60. Phelan, M. C. Deletion 22q13.3 syndrome. *Orphanet J. Rare Dis.* **3**, 14 (2008).
61. Green Snyder, L. et al. Autism spectrum disorder, developmental and psychiatric features in 16p11.2 duplication. *J. Autism Dev. Disord.* **46**, 2734–2748 (2016).
62. Horev, G. et al. Dosage-dependent phenotypes in models of 16p11.2 lesions found in autism. *Proc. Natl Acad. Sci. USA* **108**, 17076–17081 (2011).
63. Portmann, T. et al. Behavioral abnormalities and circuit defects in the basal ganglia of a mouse model of 16p11.2 deletion syndrome. *Cell Rep.* **7**, 1077–1092 (2014).
64. Sanyal, S. & Van Tol, H. H. Review the role of dopamine D4 receptors in schizophrenia and antipsychotic action. *J. Psychiatr. Res.* **31**, 219–232 (1997).
65. Rubinstein, M. et al. Mice lacking dopamine D4 receptors are supersensitive to ethanol, cocaine, and methamphetamine. *Cell* **90**, 991–1001 (1997).
66. Tan, T. et al. Stress exposure in dopamine D4 receptor knockout mice induces schizophrenia-like behaviors via disruption of GABAergic transmission. *Schizophr. Bull.* **45**, 1012–1023 (2019).

Acknowledgements

We thank X. Chen and A. Frazier for excellent technical support. This work was supported by the Nancy Lurie Marks Family Foundation and the National Institutes of Health (MH112237) to Z.Y.

Author contributions

B.R. performed the behavioral experiments and wrote the paper. K.M. performed the behavioral experiments. Z.Y. supervised the project and wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41596-020-0382-9.

Correspondence and requests for materials should be addressed to Z.Y.

Peer review information *Nature Protocols* thanks T. Takumi and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

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

Received: 11 March 2020; Accepted: 7 July 2020;
Published online: 07 September 2020

Related Links
Key references using this protocol
Qin, L. et al. *Nat. Neurosci.* **21**, 564–75 (2018): https://doi.org/10.1038/s41593-018-0110-8
Rein, B. et al *Mol. Psychiatry* (2020): https://doi.org/10.1038/s41380-020-0693-9
Duffney, L. J. et al. *Cell Rep.* **11**, 1400–1413 (2015): https://doi.org/10.1016/j.celrep.2015.04.064

nature research

Corresponding author(s): Zhen Yan

Last updated by author(s): 3-13-2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	ANY-Maze software was used for recording videos of behavioral tests and generating heat maps.
Data analysis	Data analyses were performed with GraphPad Prism 6. Experiments with two groups were analyzed statistically using two-tailed Student's t-tests (i.e. comparisons of Social Preference Index). Experiments with more than two groups were subjected to two-way ANOVA, followed by post hoc Bonferroni tests (i.e. comparisons of social vs. non-social time). Clear statistical details are provided in the "Statistics" section.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Provide your data availability statement here.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes were based on power analyses and previous experiments conducted by us (Yuen EY et al., Neuron 73:962-77, 2012; Duffney LJ et al., Cell Reports 11:1400-1413, 2015; Wei J et al., J. Neuroscience 36:2119-30, 2016) and many other labs in the field.
Data exclusions	No data was excluded from analysis.
Replication	All behavioral experiments were replicated on at least 3 independent litters. All phenotypes reported here have been reliably reproduced by multiple lab members, and many have been reported in published manuscripts from our lab and others.
Randomization	The order in which test mice were exposed to either the "3-phase S-NS" or the "2-phase S-E" method was randomized between litters to control for exposure effects.
Blinding	All manual scoring/quantification of behavioral data was performed by a genotype-blind researcher.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	The current study used a range of animals to verify the efficacy of a behavioral testing procedure. The animals used here include: male and female heterozygous Shank3+/ΔC mice expressing C-terminal (exon 21) deleted Shank3 (Jackson Labs, Bar Harbor, ME) (6-8 weeks old) and age-matched wild-type littermates; male and female dopamine D4 receptor knockout mice (D4KO) and C57Bl/6J age-matched wild-type controls (6-8 weeks); male and female 16p11.2dp/+ mice (Rein et al., Molecular Psychiatry, 2020) carrying a heterozygous duplication of the 7F3 chromosomal region homologous to human 16p11.2 along with age-matched wild-type littermates (7-9 weeks); male and female 16p11.2+/- mice (Jackson Labs, Bar Harbor, ME) carrying heterozygous deletion of the mouse chromosomal region corresponding to 16p11.2 (7-8 weeks) and age-matched wild-type littermates (C57Bl/6N129Sv); male and female Cul3flox/flox mice from Jackson Laboratory (stock #: 028349) bred with Emx1-IRES-Cre mice (stock #: 005628) to generate forebrain-specific Cul3 knockout mice (Cul3f/-) (6-8 weeks). All experiments were performed with the approval of the Institutional Animal Care and Use Committee (IACUC) of the State University of New York at Buffalo.
Wild animals	N/A
Field-collected samples	N/A
Ethics oversight	University at Buffalo IACUC

Note that full information on the approval of the study protocol must also be provided in the manuscript.