Supplementary Information

Automated long-term tracking and social behavioral phenotyping of animal colonies within a semi-natural environment

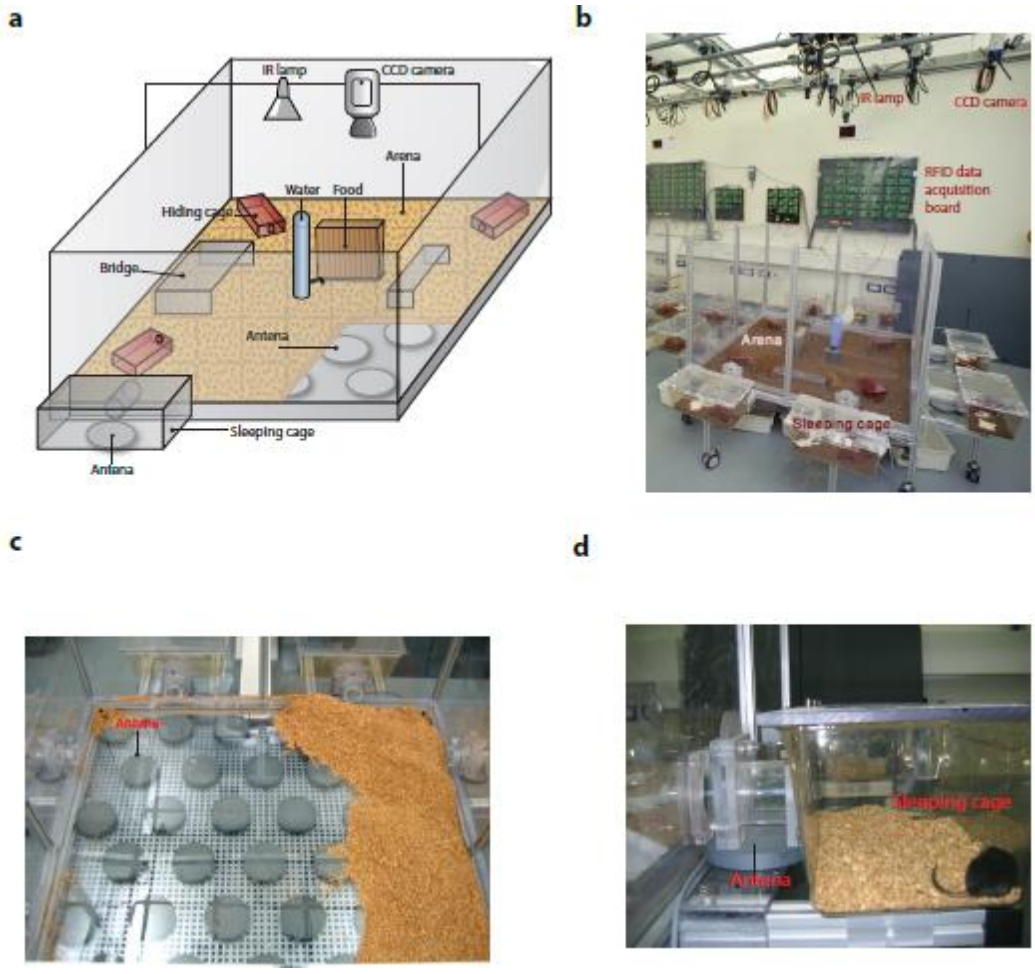
Aharon Weissbrod[1], Alexander Shapiro[1], Genadiy Vasserman[1], Liat Edry[1], Molly Dayan[1], Assif Yitzhaky[3], Libi Hertzberg[2,3], Ofer Feinerman[3] & Tali Kimchi[1*]

[1]Department of Neurobiology, Weizmann Institute of Science, Rehovot, Israel

[2]Shalvata Mental Health Center, Sackler Faculty of Medicine, Tel Aviv University, Israel
[3]Department of Physics of Complex Systems, Weizmann Institute of Science, Rehovot, Israel

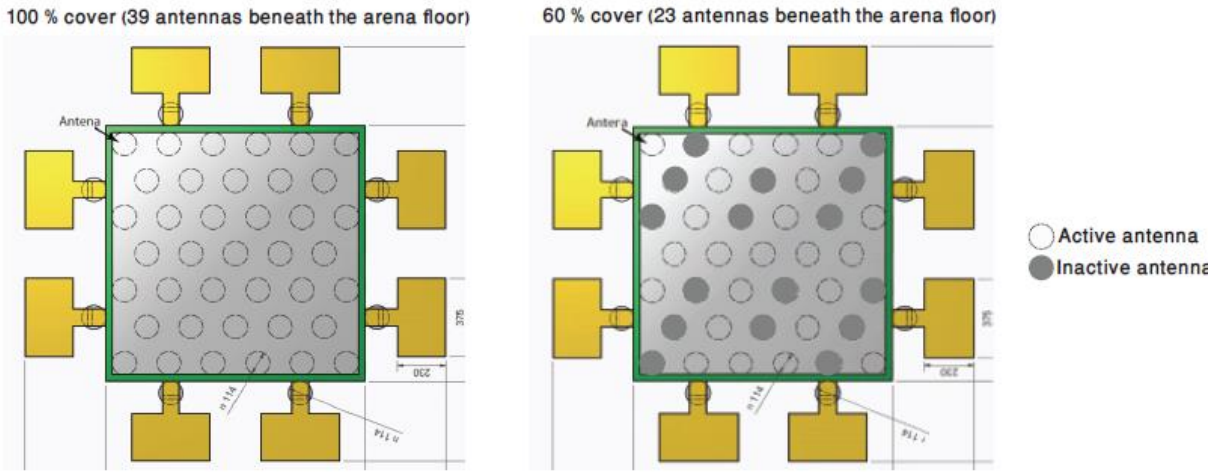
[*]Correspondence should be addressed to: Tali Kimchi (tali.kimchi@weizmann.ac.il)

Supplementary Figures



Supplementary Figure S1: Automated tracking system for colonies within a large semi-natural setting. (**a-b**) Side views of the testing apparatus. The enclosure comprises a large exploratory arena with eight peripheral sleeping cages that are freely accessible to the mice. (**a**) The arena floor is evenly covered with 2–3cm of sawdust bedding and is equipped with shelter boxes, bridges, a free access feeder filled with rodent pellets, and a water bottle. Mice behavior is recorded by a video-tracking system that comprises (**a**) ceiling-mounted CCD camera, (**b**) infra-red illuminators, and (**c–d**) a radio frequency (RFID) tracking system, which consists of an array of RFID antennas positioned beneath the floor of the exploratory arena and beneath the short transparent tubes connecting the sleeping cages to the arena.

a



b

System performance in tracking position and identity of three interacting mice with respect to the number of active RFID antennas

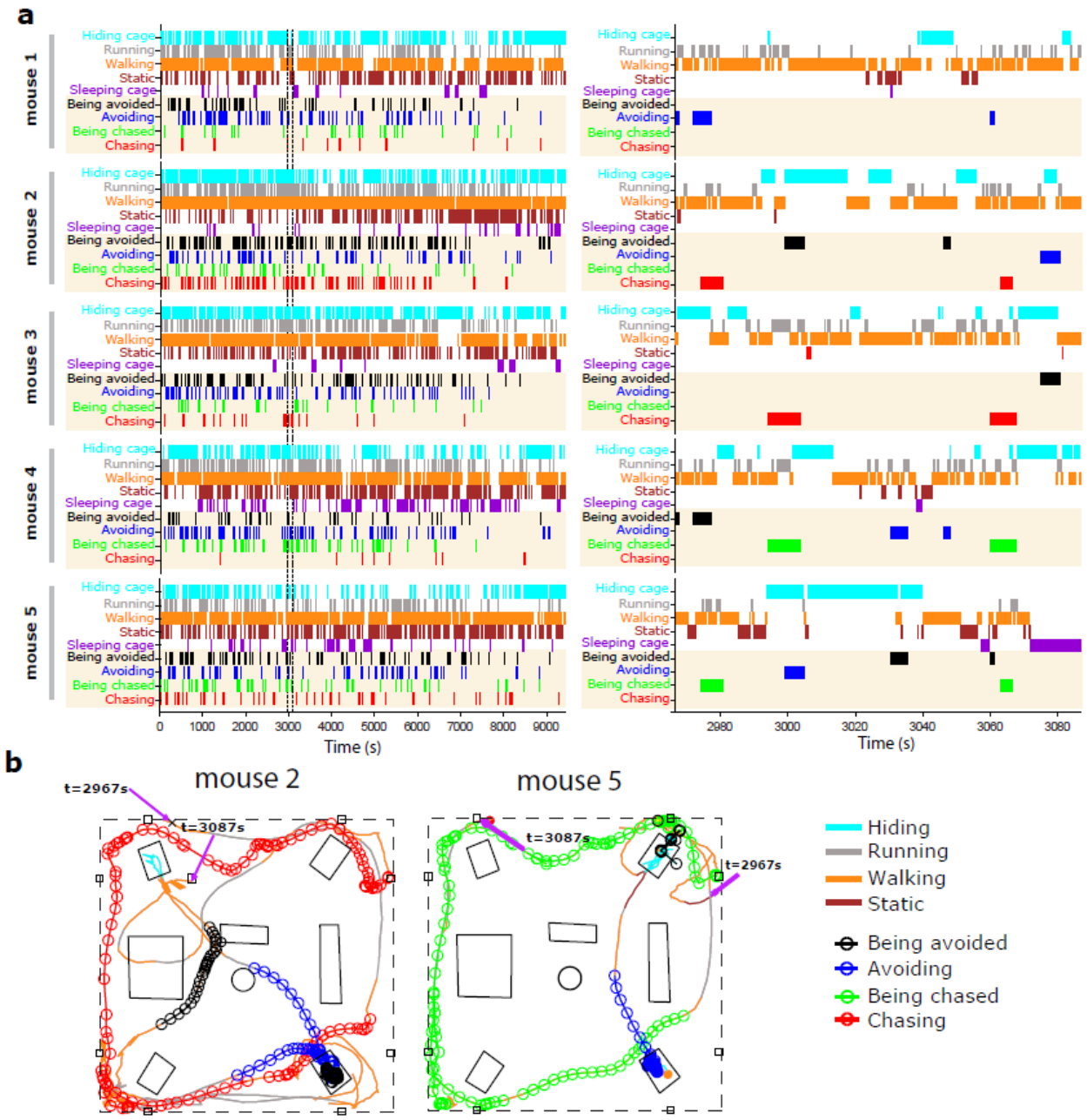
Number of antennas active (%)	Median tracked mice	Swaps identities (%)	Unknown identities (%)	Overall system performance (%)
39 (100)	3	0.11	0.17	99.72
35 (90)	3	0.11	0.25	99.64
31 (80)	3	0.46	0.31	99.23
23 (60)	3	0.99	0.20	98.81
8 (20)	2	8.11	0.27	91.62

c

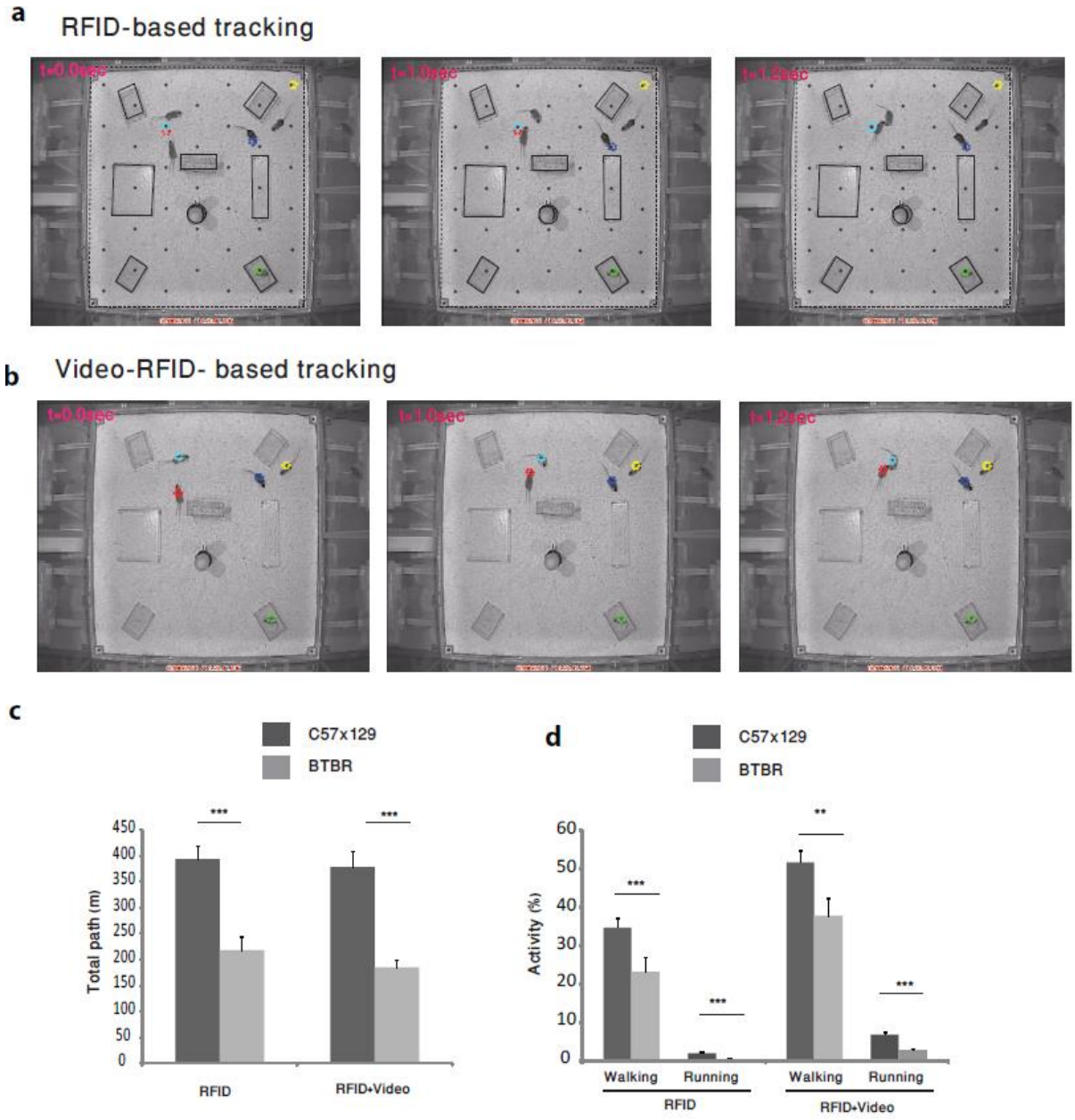
System performance in tracking position and identity of six interacting mice with respect to the number of active RFID antennas

Number of antennas active (%)	Median tracked mice	Swaps identities (%)	Unknown identities (%)	Overall system performance (%)
39 (100)	6	2.63	1.63	95.74
35 (90)	6	2.74	2.84	94.42
31 (80)	6	2.93	4.01	93.06
23 (60)	5	1.58	7.86	90.56
8 (20)	4	0.89	17.58	81.56

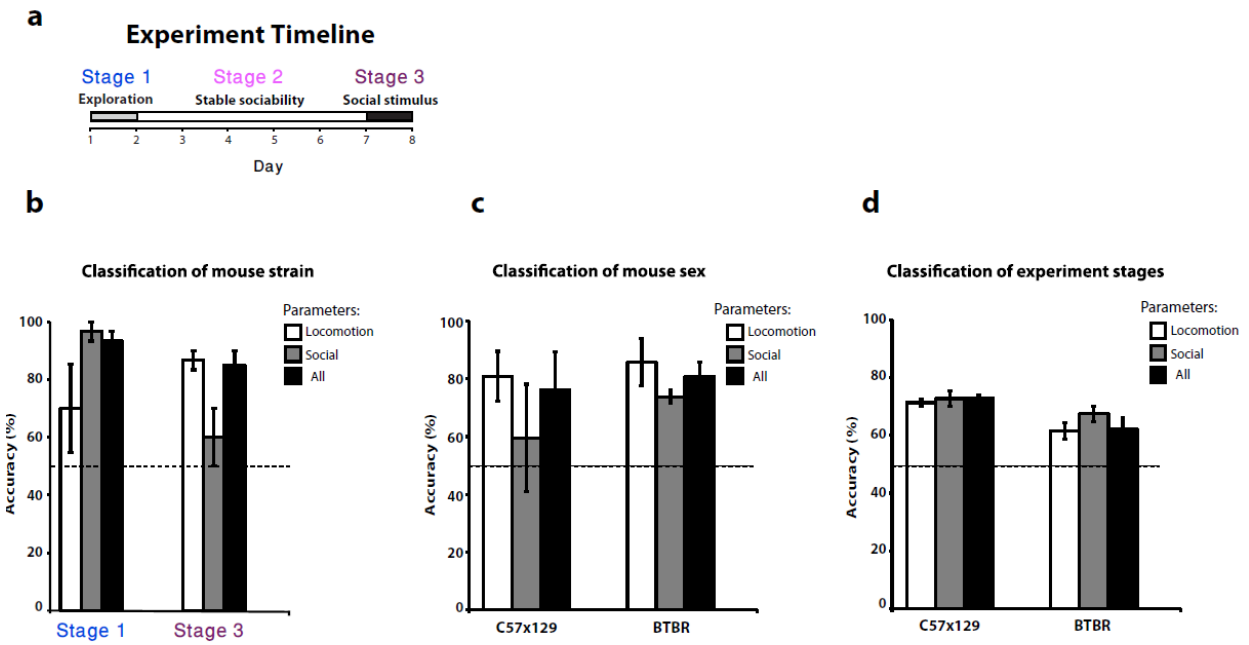
Supplementary Figure S2: System performance as a function of the density of the active RFID antennas. (**a**) Position of the RFID antenna array positioned beneath the arena floor. (**b**) System performance in tracking the spatial position and identity of group of six mice with respect to the density (number) of active RFID antennas.



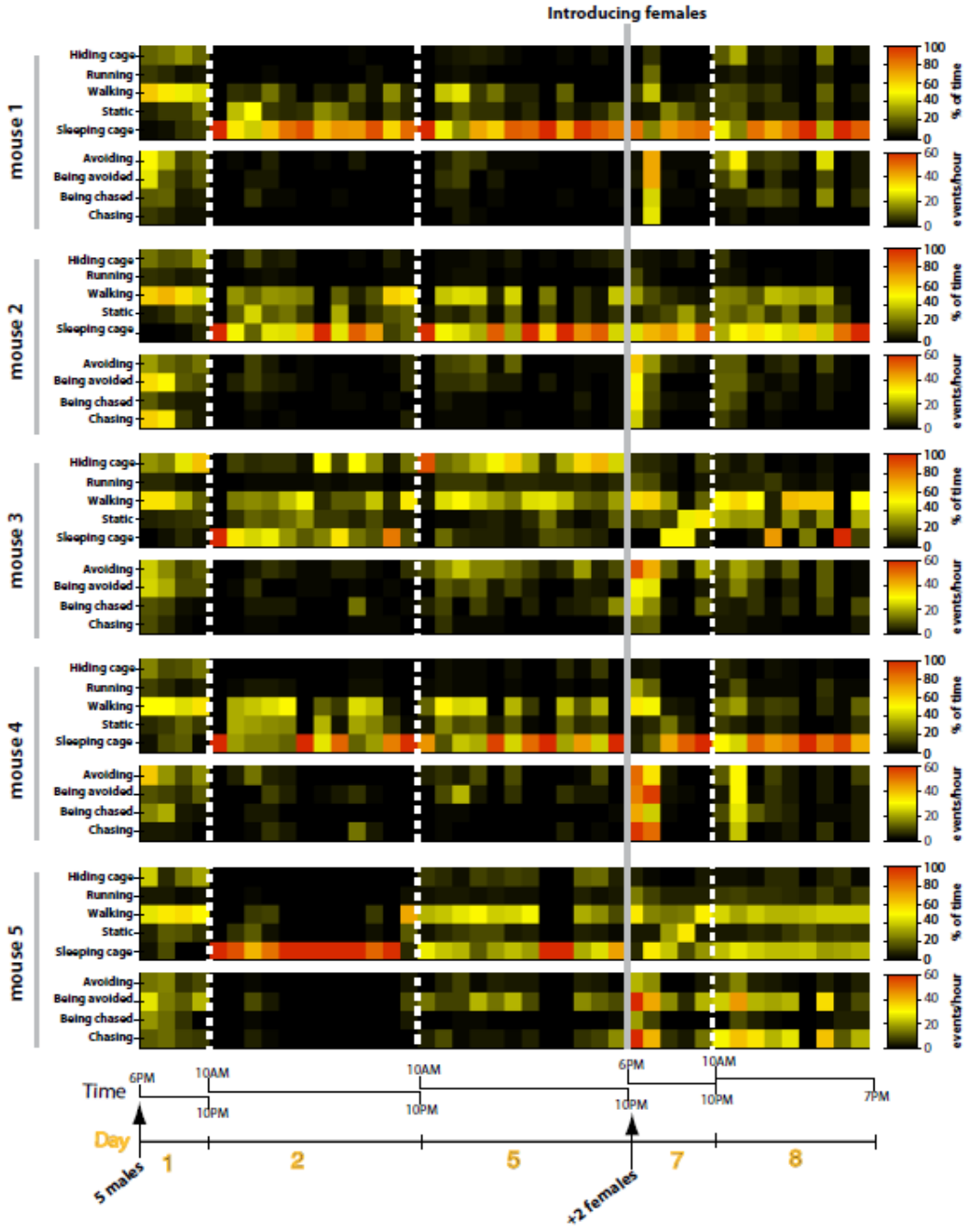
Supplementary Figure S3: Example of ethograms of five male mice showing a subset of the behavioral features scored by the system in a single day. (**a**) Behavioral ethogram including five locomotion features and four social behavior features scored over 120 minutes (left panel) and two minutes (right panel). (**b**) An example of a two minute ethogram of the nine socio-behavioral features examined overlaid on the trajectory of two tested mice (the same mice as presented in panel (**a**)). The beginning and the end of each trajectory are marked with an arrow.



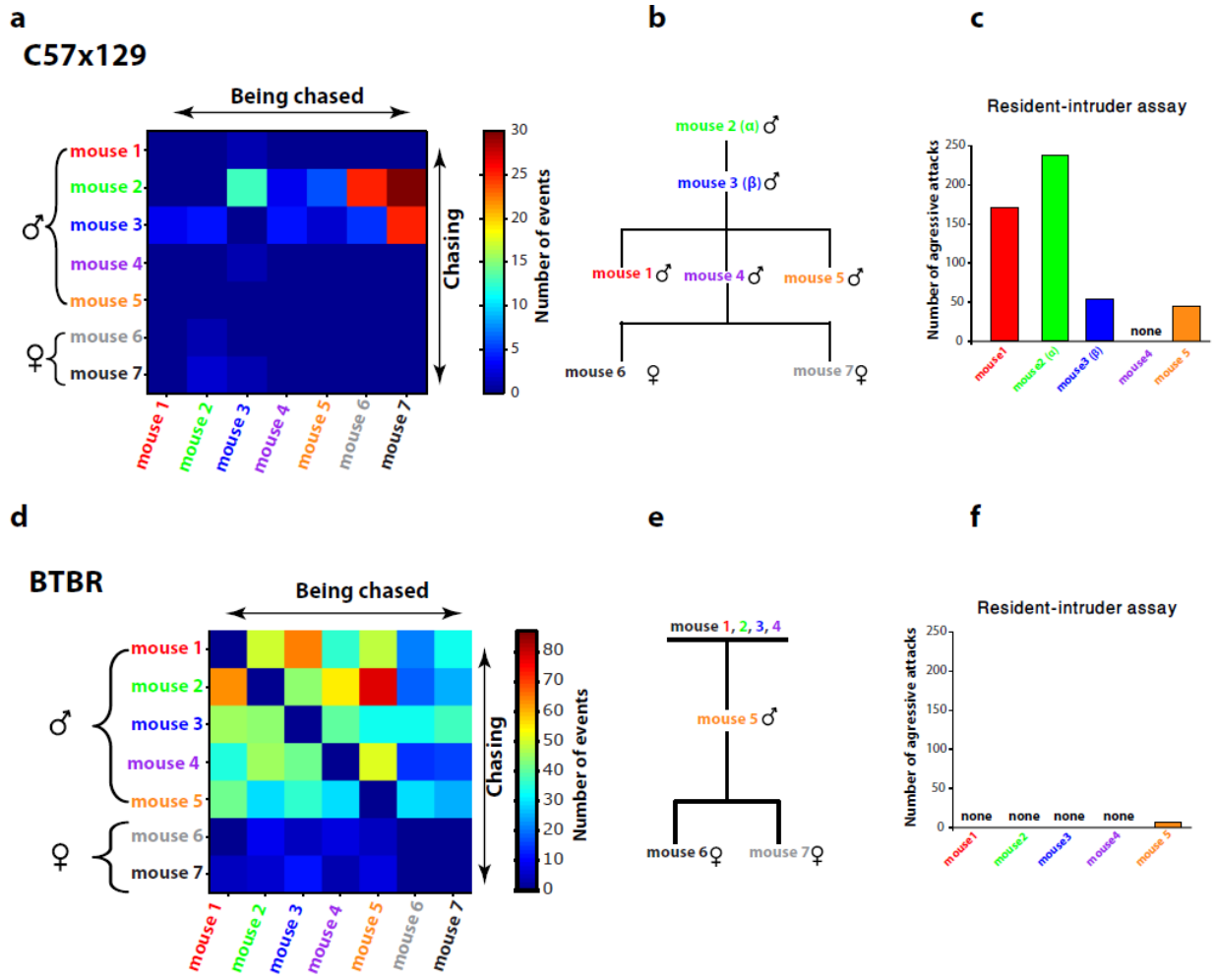
Supplementary Figure S4: Tracking and locomotion phenotyping based on the RFID tracking system versus the Video-RFID tracking system. (**a-b**) Consecutive video snap shots of localization of position and identity of five mice using (**a**) RFID tracking system and (**b**) video-RFID fused tracking system. (**c**) Travelled path and (**d**) percentage of walking and running time per one hour of testing (day 1) for two strains (n=15 individuals in each strain), as classified using only the RFID tracking system data or the video-RFID tracking system data. Data are presented as the mean ± SE; *** $P < 0.001$ between strains, ANOVA tests.



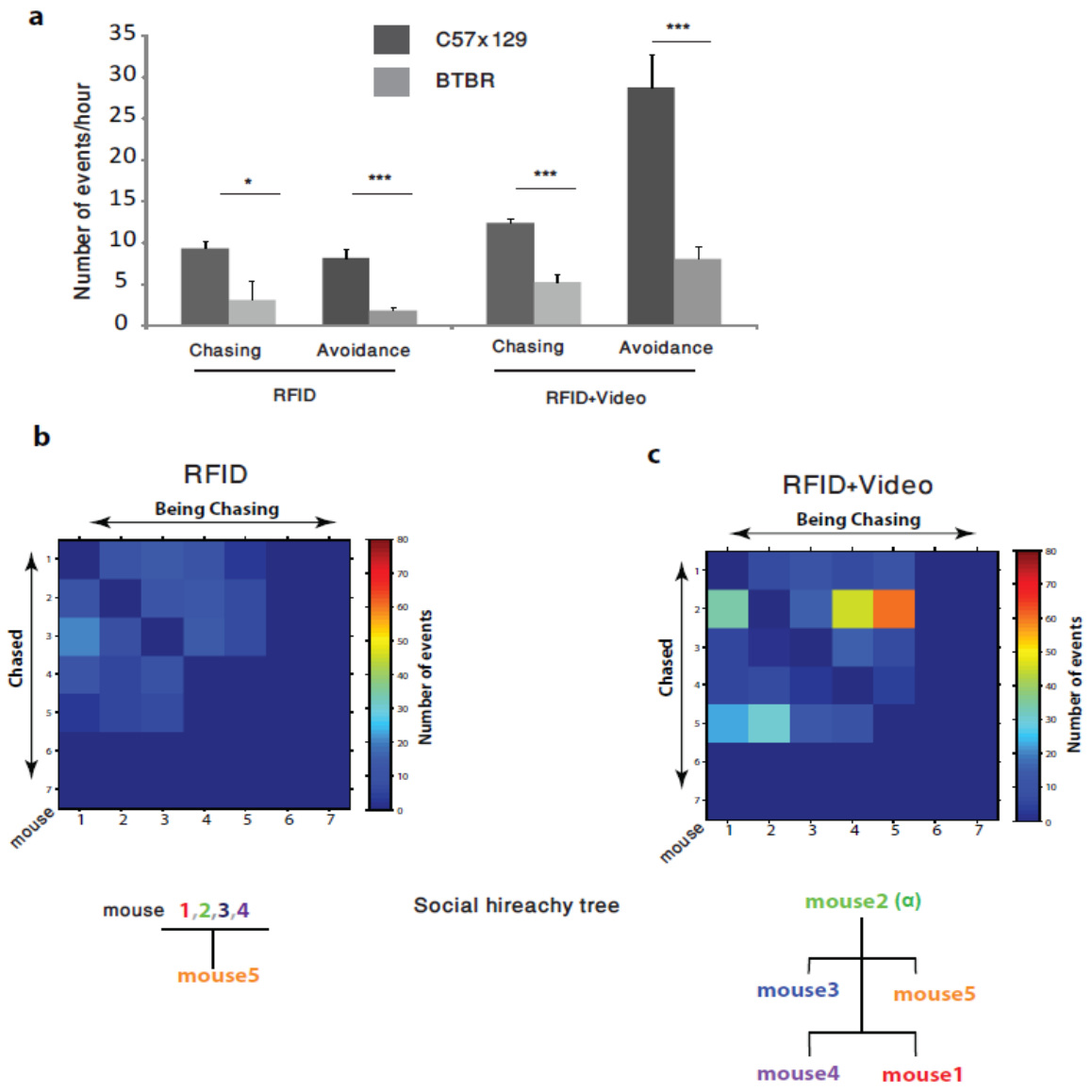
Supplementary Figure S5: Accuracy of strain, sex, and behavioral state prediction classification from automatically quantification set of locomotion and social parameters. (**a**) Analysis of the behavioral dynamic throughout the experimental days revealed that the experiment can be subdivided into three distinct stages. Stage 1 (exploration, first day in the enclosure) — all individuals engage in a high level of social interactions. Stage 2 (social behavior stability, days 2–6) — mice spend significantly less time in daily social interactions compared with stage 1. Stage 3 (social stimulus, days 7–8) — a dramatic rise in social interactions triggered by the introduction of two alien females into the arena. (**b–d**) We found that, using either the locomotion parameters (five behavioral features: sleeping, hiding, static, walking, running), social parameters (four behavioral features: chasing/being chased, avoiding/being avoided), or all parameters, we can accurately classify the mouse strain (**b**) its sex (**c**) and the day in the experiment (experimental stage) from which the data were obtained (**d**). Data accuracy is presented as the mean (± SE) (see Supplementary Methods).



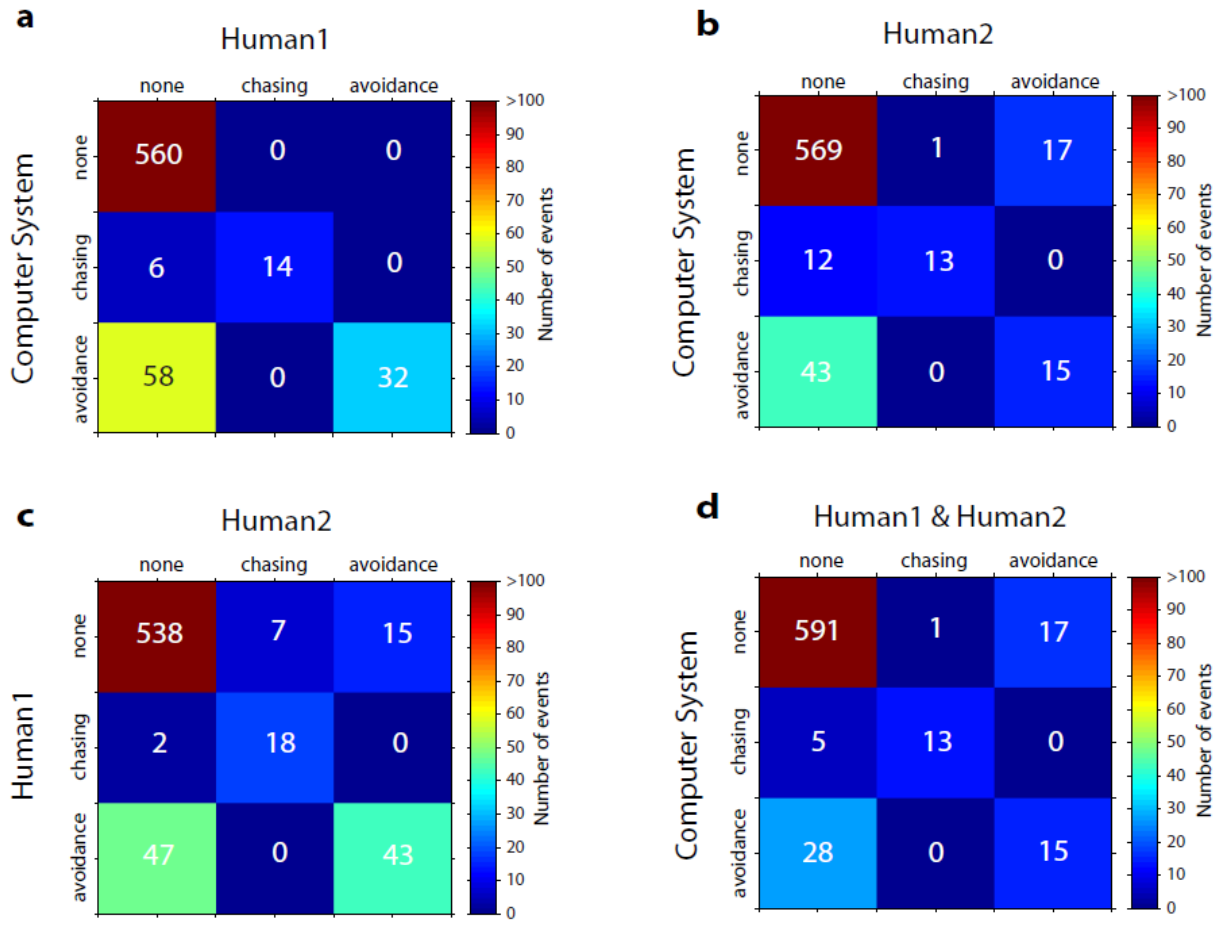
Supplementary Figure S6: Example of ethograms of five male mice showing a subset of the behavioral features scored by the system over eight days. Each column represents the frequency of the behavior during a one hour period and each row represents a behavioral parameter. The experiment was performed under reverse light cycling (10 am white light off, 10pm white light on) and behavioral analysis was focused on the dark period during which the mice are more active.



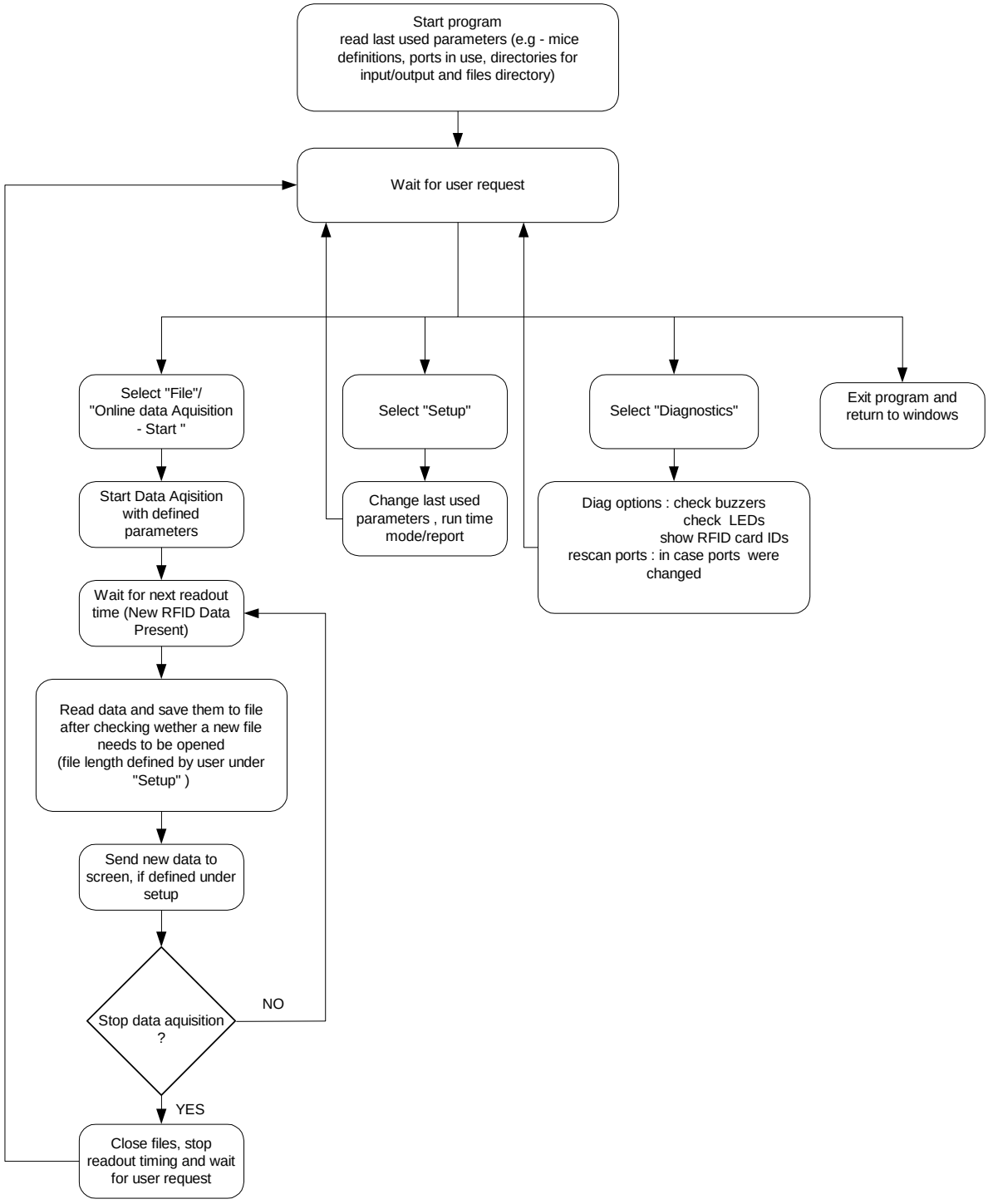
Supplementary Figure S7: Dominant-subordinate social hierarchy is automatically quantified by our system. (**a, d**) Chasing and being chased interaction matrix for C57x129 (a) and BTBR (**d**) groups of mice. (**b, e**) Social hierarchy trees based on the chasing and being chased interaction matrices for C57x129 (**b**) and BTBR (**e**) groups of mice on day 7 of the experiment. Overall, the C57x129 group presented a social hierarchy with a single identified dominant male (mouse #2) and sub-dominant male (mouse #3). In contrast, the BTBR group presented no social hierarchical organization and a dominant male was not established (see also Fig. 3). (**c, f**) Results of classical resident-intruder assays on the C57x129 (**c**) and BTBR (**f**) mouse groups previously tested in our system. Data present human-quantification of the number of aggressive attacks presented by each resident male mouse (n = 5) toward the intruder male mouse during a 15 minute session.



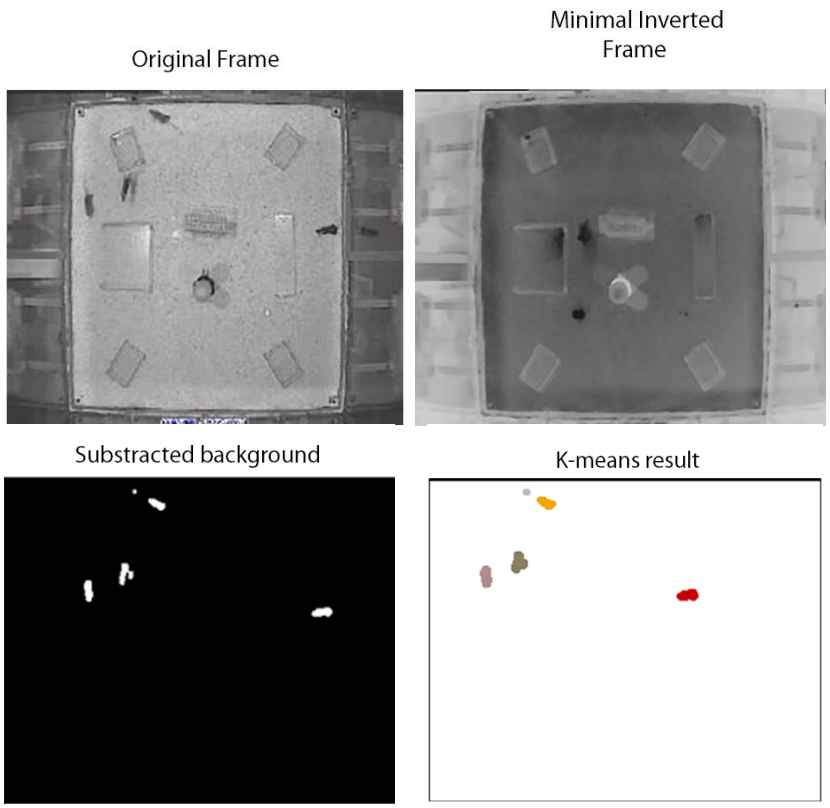
Supplementary Figure S8: Social behavior phenotyping based on the RFID tracking system versus the Video-RFID tracking system. (**a**) Chasing and avoidance behaviors for two strains (n=15 individuals in each strain), as classified using only the RFID tracking system data or the video-RFID tracking system data. Data are presented as the mean ± SE; * $P < 0.05$, *** $P < 0.001$ between strains, ANOVA tests. (**b-c**) "Chasing" and "Being-chased" interaction matrix of a C57x129 group, as classified using only the RFID tracking system data or the video-RFID tracking system data. The matrix was used to define the social ranking of each individual in the group. Note: Social hierarchy organization could be defined only by using the video-RFID tracking system.



Supplementary Figure S9: Accuracy of automatic classification of social behavioral features. (**a**) Social behavior labeling confusion matrix of human-1 observer versus machine classifier. (**b**) Social behavior labeling confusion matrix of human-2 observer versus machine classifier. (**c**) Confusion matrix of human-1 observer versus human-2 observer for the classification of social behaviors. (**d**) Social behavior labeling confusion matrix of human-1 and human-2 (only events that were with agreement between the two humans) observer agreements versus the machine classifier (see Supplementary Methods). Data in the matrices indicates the number of events classified as no social interaction ("none"), chasing behavior ("chasing") or avoidance behavior ("avoidance") by the human observer and the machine classifier.



Supplementary Figure S10: RFID-based data acquisition software. Flowchart detailing the basic user-defined settings of the software and the operation steps to initiate and terminate RFID data acquisition (see Supplementary Movie 7 and http://www.weizmann.ac.il/neurobiology/labs/kimchi/content/downloads).



Supplementary Figure S11: Defining the X-Y coordinates of each mouse in the arena based on the collected video data. For each video frame recorded, the algorithm detected each mouse in the arena and determined its coordinates by running several sequential image analysis steps, as explained above. Image analysis included definition and subtraction of the overall background from each frame, followed by definition and clustering the pixels of each individual mouse (using k-means clustering) to define the location of its center of mass (http://www.weizmann.ac.il/neurobiology/labs/kimchi/content/downloads).

Supplementary Methods

I. Enclosure (arena) set-up

The enclosure was comprised of a central square exploratory arena measuring 119.2 × 119.2 × 80 cm (L × W × H) and constructed of transparent polycarbonate boards attached to each other by aluminum square rods (KANYA LTD). A white opaque plastic sheet was pasted over the arena's floor to avoid light reflection from the illumination positioned above the arena.

Two standard mouse cages (15 × 23 × 15 cm) were attached to each of the four arena walls by short Perspex tubes to serve as sleeping nests. Mice were allowed to freely move between the exploratory region and the sleeping nests (**Supplementary Fig. S1**). The system was also furnished with: (1) a custom-designed radiofrequency-identified (RFID) position-tracking system and (2) an array of ceiling-mounted low-light sensitive video cameras (note: only a single camera above the center of the exploratory arena is required for system operation) connected to a multi-channel digital video-recording system and an array of infra-red illuminators to enable observation of mouse behavior during the dark period.

II. RFID-based setup

1) RFID antenna array: The antenna array included: 39 coil antennas positioned beneath the exploratory area, for the purpose of mouse identification and tracking in the arena area, and another 8 antennas located under each of the tubes connecting the arena with the nesting cells, to detect mouse movements into and out of the nesting cells. All 47 antennas in the system were identical, being round in shape, with a diameter of 110 mm, an inductance of approximately 380 μHy, and encapsulated in a PVC case. The distance between the antennas was set at 195 mm, center to center (**Supplementary Fig. S1**).

2) RFID readers (decoders): TM613 single coil antenna driver boards were piggyback-mounted onto LID650 decoders (Dorset (Trovan) Electronic Identification Systems). The decoders were powered by a 12 VDC linear power supply in order to avoid the power supply disturbing the decoders. The reading rate per decoder was ~40 reads/second for a single

transmitting transponder. An asynchronous 2-wire communication protocol (RS485 standard) at a baud rate of 57,600 was chosen. At this baud rate, it is necessary to limit the number of ports per net in order to optimize the reading rate. For the whole system, we used 4 nets and received an average of ~150 readings/second. All the decoders were time-synchronized in order to avoid beat frequency interference between the decoders.

3) RFID power supply: A modular M 15–16 HE power supply (Delta Elektronika BV) was used. A single decoder requires a maximum current of ~200 mA, so the entire system required up to 9.4 A.

4) RFID-based data acquisition and storage hardware: The RFID-based data were streamed to a standard PC computer equipped with a National Instruments NI PCIe-8431/8 RS485/RS422 8-port asynchronous serial interface card.

5) RFID data acquisition software: The data acquisition software was written in Lab Windows/CVI VER 10 and functioned on a data polling method at a user predefined rate. At startup, the software searched all the possible data-transmission ports that were available for communication with the decoders and marked them on the graphical user interface (GUI). Once the search was complete, all the available ports appeared on the GUI window along with their individual number, which allowed visual presentation of the time at which each port detected or did not detect data from the transponders (**Supplementary Movie 7** and **http://www.weizmann.ac.il/neurobiology/labs/kimchi/content/downloads**).

III. Video-based recording system

Commercial CCD cameras (>570TLV black/white; MTV-13V5HC, Mintron) were mounted above the enclosure, as follows: one camera was positioned exactly above the center of the exploratory region and was the only video camera essential for the system's operation. An array of infra-red illuminators (FMVL-R50, MSDV) was mounted above the enclosure to evenly illuminate it during the dark period. The array of cameras was connected to a commercially available digital recording unit (DVR; CTTI, Ltd) that allows simultaneous recording of the video streaming from up to 16 channels and at 30 frames per second. The DVR recording computer was synchronized with the PC computers that were simultaneously recording the RFID data. Time-synchronization is essential for the system's operation and was accomplished using the Absolute Time Server software (FlexibleSoft) that was installed on both the video and RFID computers. The DVR computer was set to serve as the "time server" (the time setter of the experiment) and the RFID recording computer was configured to serve as the "time client" computer.

IV. Video and RFID data fusion and mouse position tracking software

Fusion of video and RFID data was implemented by a software application written in OpenCV C++. The final output of the software was a text list of the X,Y coordinates of each mouse in the enclosure. Data were extracted and stored in a file ("csv" format) for further behavioral phenotyping analysis (see **http://www.weizmann.ac.il/neurobiology/labs/kimchi/content/downloads**).

To extract the position of the mice from the integrated video and RFID-based tracking data for each video frame recorded, the following three main steps were performed: (1) The position of each tested mouse was detected in the arena coordinate plane based on video tracking with high spatiotemporal resolution. (2) Each mouse was unambiguously identified and its position was estimated at a low spatial resolution based on RFID data. (3) Data streams (1) and (2) were fused to precisely track mice at a high resolution. To achieve such resolution, the following data processing was required:

1) Calculating the values of the background: We built a background mask, B, for each video frame recorded, where $B = b_{ij}$ ($1 \le i \le m$ for frame height, $1 \le j \le n$ for frame width) by estimating the minimum $F_{i,j}^{min} = \min_t \left(255 - \left\{F_{i,j}^t\right\}\right)$ and maximum $F_{i,j}^{max} = \max_t \left(255 - \left\{F_{i,j}^t\right\}\right)$ pixel values of the inverted image frames for all the video frames F_t over time t. Since, in our experiments, mouse pixels had lower values than the background, $F_{i,j}^{\frac{max}{min}}$ values were effectively used to set the threshold for (negative) background removal.
2) Separating the image pixels into mice (foreground objects) and background pixels: We defined the threshold as the average image intensity $Tr_{i,j} = \frac{F_{i,j}^{max} + F_{i,j}^{min}}{2}$ and removed the background by setting all the pixels that were below the threshold to zero: $\widetilde{F_{ij}^t} = 0 \; if \; \widetilde{F_{ij}^t} < Tr_{i,j}$. As a result, the background pixels have a value of 0 and all the mouse pixels are equal to 1.
3) Classification of pixels to individual mice: In order to estimate the center of mass of each mouse in the arena, we performed an analysis to cluster the pixels belonging to each mouse using the k-means clustering method[45]. This resulted in a set of mouse positions $\{P_t^k : (x, y) \in \mathbf{R^2}\}$ for each frame t and mouse k. When two or more mice were in physical contact with each other, the k-means algorithm was not sufficient to segregate their pixels into different clusters. We solved these cases by estimating the area $S_k \; k \le N$ (N, number of mice) of every cluster. When the area S_k was greater than some $\hat{S}$ pixels limit, we assumed that there was more than a single mouse in that position. The total number of mice in the arena at a specific frame t was estimated using the RFID readings.

4) Finding the RFID-based coordinates for each video frame: There was a need to estimate RFID-based positions for video frames which lacked RFID transponder numbers. This was performed by averaging the positions (P) of several RFID readings and then normalizing

each one as a function of the time difference from the estimated time frame t_0: $\overline{P_{RFID}}(t_0) = \sum_t P_{RFID}(t) e^{-\frac{\Delta t^2}{\sigma^2}}$, where $P: (x, y) \in \mathbf{R}^2$, $\Delta t = (t - t_0)$, and $\sigma = 0.03\ s$.

5) Building a mask function to filter out video-RFID pairs with low probability: We set the mask function, as follows:

$$M\left(P^k_{video}(t)\right) = 0 \text{ if } R = \frac{\Delta t + 0.03}{0.06666} V_{max} \times 180 < \left|P^k_{video}(t) - P^s_{RFID}\right|$$

Otherwise: $M\left(P^k_{video}(t)\right) = 1$ (here: mask M depends on 3 parameters – (t, k, s)).

$V_{max} \cong 400\ mm/s$ is the maximal mouse velocity, and *k* and *s* are the video and RFID tracking indexes respectively; $1 \leq k \leq 7$, $1 \leq s \leq 7$. These values can be redefined by users, if required.

6) Video-RFID position anchor points: Since the RFID data have a relatively low spatial resolution and are acquired only when the mouse is in proximity to the RFID antennas, we predicted mouse trajectories based on the points in time when we were sufficiently certain of their spatial location. We refer to these points as anchor points, A^k_t, for each mouse k and for a specific time frame t. When the distance between the mouse spatial location based on the video and the RFID tracking systems was smaller than the maximal distance between adjacent antennas, the point was labeled an anchor point.

7) Finding mouse trajectories between anchor points: The algorithm that predicts mouse trajectories between anchor points was based on finding the minimal distance trajectory from all the given possibilities. The trajectories were found by minimizing the function:

$$D = \sum_{i=T_a}^{T_{a+1}} \sum_{k,s} \frac{\left|P^k_{t_i+1} - P^k_{t_i}\right|^2 + \left|P^s_{t_i+1} - P^s_{t_i}\right|^2}{\sigma_{k,s}^{t_{i+1}\,2}}$$

where T_a is the frame index of the anchor point, k and s are indices of different positions P at time t between successive frames, and $\sigma_{k,s} = \frac{|P^k - P^s|}{U}$, where U is the maximal distance between two points in the arena (*i.e.*, the length of the diagonal).

I. Mouse social behavior extraction based on mouse trajectories

Socio-behavioral patterns between pairs of mice (Mouse A and Mouse B) where characterized using the following behavioral features (**Supplementary Fig. S11** and **Supplementary Movies 3-6**): chasing/being chased and avoiding/being avoided.

1) Chasing behavior: Chasing behavior was detected for each pair of mice. First, the time frame of the experiment was partitioned into segments that were assumed to include, at most, one chasing event, based on the distance between each tested pair of mice. This was based on the assumption that during the chase, mice are positioned relatively close to each other and are further apart before and after the chase. Thus, only the segments in which the distance between the mice was less than a specific threshold were analyzed as chasing events. Then, we assumed that during the chase, both mice are moving in the arena, therefore only the segments in which the path traveled by each mouse was greater than a specific distance were analyzed further. The next condition for positive chasing behavior was a high correlation between the trajectories of the two mice. Only the segments with an X-Y position correlation above 0.7 were classified as positive for chasing. The mouse that was engaged in chasing behavior was classified according to the directions of the unit vectors $\overrightarrow{v_1}$ (the unit vector of mouse A moving from position (x_{t-1}^1, y_{t-1}^1) in frame $t-1$ to position (x_t^1, y_t^1) in frame t) and $\overrightarrow{v_2}$ (the unit vector between the positions of mouse B at (x_t^2, y_t^2) and mouse A at (x_t^1, y_t^1). If vectors $\overrightarrow{v_1}$ and $\overrightarrow{v_2}$ pointed in the same direction, then mouse B was defined as chasing mouse A, otherwise mouse A was defined as chasing mouse B.

2) Avoidance behavior: Avoidance behavior was tested for each pair of mice. As with the detection of chasing behavior, we assumed that during avoidance behavior, for at least several successive frames, the two mice will be in close proximity (*i.e.*, less than a specific distance from each other) and that one of the interacting individuals (the avoided mouse) will accelerate sharply (above 120 cm/sec) for a few sequential video frames. The avoiding mouse was therefore classified according to: 1) the direction of unit vector $\overrightarrow{v_1}$ between mouse A at position (x_{t-1}^1, y_{t-1}^1) in frame t-1 and position (x_t^1, y_t^1) in frame t; and 2) the

direction of unit vector $\overrightarrow{v_2}$ between mouse B at position (x_t^2, y_t^2) and mouse A at position (x_t^1, y_t^1) during the escape phase.

II. Validation of social behavioral parameters

Two well-trained human observers manually quantified 670 short video clips in which at least two mice (each color-labeled by our automated tracking system) were present in the arena with at least one individual in motion. The two observers were requested to independently classify the behavior in each of the video clips into one of the following: mouse A chasing mouse B, mouse A being chased by mouse B, mouse A avoiding mouse B, mouse A being avoided by mouse B, or the mice did not socially interact. The behavioral quantifications of the human observers were compared with each other and with the computer-based quantification provided by our system.

III. Social hierarchy tree

The social hierarchy tree was built based on the chasing/being chased pairwise interaction data[44]. First, the dominance, equality, or undetermined relation was set for each pair of mice. The decision was based on determining the ratio of "wins" (in chasing events) with respect to the total number of chasing/being chased interactions. The experimental result was tested using a binomial probability model (binomial cumulative density function). Dominance was determined if $\frac{cdf_{0.7}^{1}}{cdf_{0.4}^{0.6}} > 1.5$, where $cdf_{p_1}^{p_2} \sim \binom{n}{k} \int_{p_1}^{p_2} t^k (1-t)^{n-k} dt$, where n is the total number of interactions and k is the number of "wins" by the more dominant of the two mice. The relationship between the mice was considered equal if the inverse relationship, $\frac{cdf_{0.4}^{0.6}}{cdf_{0.7}^{1}} > 1.5$, and otherwise it was impossible to determine the exact relationship between the mice (**Fig. 3**).

XVII. Classification of mouse strain, sex, and day of experiment

We have used the Naïve Bayes method (Matlab, Statistics Toolbox) to classify three types of features: 1. mouse strain (C57x129 or BTBR), 2. mouse sex (male or female) and 3. day of the experiment (stage 1, 2 or 3). For this purpose, we used a dataset of six independent experiments (three experimental groups per mouse strain), which included C57x129 males (n = 15), BTBR

males (n = 15) and C57x129 females (n = 12). For classification of mouse strain (C57x129 or BTBR) we used two experiments from each strain as training runs and the remaining independent experiments for testing. We ran three combinations of experiments for training and testing based on data gathered in stage 1 (day 1, exploration stage) and stage 3 (days 7–8, social stimulus stage) and then calculated the mean result. Similar assays were used to classify mouse sex (male or female), except that all data were taken from stage 3 of the assay. For classifying the day of the experiment (stage 1–3), we used two experiments from each strain as training runs and the remaining independent experiments for testing. Both training and testing rely on datasets gathered on a single day.

Supplementary References

44. Izar, P., R. G. Ferreira, et al. (2006). "Describing the organization of dominance relationships by dominance-directed tree method." Am J Primatol **68**(2): 189-207.
45. Kanungo, T., D. M. Mount, et al. (2002). "An efficient k-means clustering algorithm: analysis and implementation." Pattern Analysis and Machine Intelligence, IEEE Transactions on **24**(7): 881-892.