**Role of central serotonin in social life of laboratory mice in a naturalistic environment**

Weakened social knowledge and competency in serotonin-deficient mice in naturalistic environment

Marion Rivalan a \*, Valentina Mosienko b, c, Lucille Alonso a, Patrik Bey a, d, Alexia Hyde a, Michael Bader b, York Winter a §, Natasha Alenina b §

a Humboldt University Institute of Biology, Chair of Cognitive Neurobiology, Philippstr. 13, 10099 Berlin, Germany.

b Max-Delbrueck-Center for Molecular Medicine, Robert-Roessle-str.10, 13125 Berlin, Germany.

c University of Exeter, College of Medicine and Health, Prince of Wales Road, EX4 4PS, Exeter, UK

d Brain Simulation Section, Department of Neurology, Charité Universitätsmedizin Berlin, Germany.

§ Authors contributed equally to this work.

\* Correspondence and present address:

Marion Rivalan, Charité Universitätsmedizin Berlin, Exzellenzcluster NeuroCure, CharitéCrossOver, Animal Outcome Core Facility, Virchowweg 6, 10117 Berlin, Germany

e-mail: marion.rivalan@charite.de

tel.: +49 (0)30 450 539 743

**Potential journals**

Frontiers in Molecular Neuroscience, Research Topic: Neuroscience of Serotonin – New Vistas

(IF 3.72) – *deadline end of March*

Translational psychiatry (5.6) - *very short introduction.*

*https://www.nature.com/documents/tp-gta.pdf*

Scientific reports (4.12)

Biological psychiatry (10+)

**Rest to do list**

- Add the citations +++

- check the genotype of mice (NA)

- suggest reorganization results of Table 2 (NA)

- complete “ethic statement” (NA)

- give MR feedback on text (all)

- homogenise the legend of the figures ({R..})

- homogenise citation of figures/tables in the MM

- cite figures: Fig1.A or Fig.1A?

- write Abstract

- “At six weeks of age, four mice of the same genotype were transferred to the same cage”

<- NA: from different litters/cages?

- Fig2 1st hour KO rats – check data (MR/LA)

**Acknowledgements**

the Northcott Devon Medical Trust Foundation Grant to VM (TB/MG/NO5002)

**Abstract**

**Introduction**

Aggression is an adaptive behaviour, often the result of a competition (Nelson and Trainor 2007 – see Kiser 2012). It is an important social behaviour that ensures the stability and the prosperity of a social group (see DFG grant: van loo 2003, Sapolsky 2005, Wang 2011). When it cannot be avoided, an aggression is typically short (Kiser 2012) and meant to achieve a purpose such as to acquire or keep resources (i.e. territory, mating partners, food…). On the pathological side escalated aggression or violence, excessively and repeatedly hurt others or the perpetrator itself, happens in all contexts and is devoid of a communication purpose (de Boer 2009). This extreme behaviour is detrimental to the individual (acute death, invalidity…) and severely disrupts the group structure, its security and comfort (WHO). In humans, interpersonal violence has a high economical cost (WHO) and a better understanding of how aggression impacts the group’s structure is critical for “diagnosis, prevention, and treatment, but also for guidance of public and judicial policies” (Miczek 2007).

Uncontrolled aggression and violence are diagnostic criterions of different psychiatric disorders, such as schizophrenia, alcoholism, intermittent explosive disorder, autism or dementia (Lesch et al. 2012, see Volavka 2012 in Niederkofler 2016). Considering a dimensional and trans-diagnostic view of mental disorders, the comorbidity of uncontrolled aggression across different psychopathologies makes it a good potential endophenotype of mental disorders [Niederkofler 2016 + RDoC] with the neurobiological (and heritable) substrates of pathological aggression being the same across different psychiatric disorders (DSM-V, Gottesman 2003, Gould 2006, Robbins 2012 – see grant). The identification of the neurobiological substrates of aggression is thus essential to understand the processes leading to complex brain disorders (Kalueff et al., 2015) and its potential global impact on society.

Serotonin (5-hydroxytryptamine, 5-HT) is a monoamine that plays a critical role in the regulation of impulsive and aggressive behaviours in humans and animals (see multiple references in Niederkofler 2016). Mice with a congenital lack of serotonin in the brain due to the lack of rate-limiting serotonin-producing enzyme TPH2 (Tph2-deficient mice, Alenina 2009) were found to be very belligerent animals. They show higher levels of aggression toward strangers, poor social recognition abilities and an impulsive like phenotype but also presented behavioural abnormalities similar to human symptoms of Autistic Syndrome Disorder (ASD) and Impulsive related disorders (Mosienko 2012, Kane 2012, Beis 2015, Angoa-Perez 2012, Kästner 2019). In line with the search for trans-nosological symptoms of mental and neurological disorders, the Tph2-deficient mice appear to be an interesting model of pathological aggression modulated by the serotonergic central system.

The Visible Burrow System (VBS) is a semi natural habitat first developed in rats (Blanchard and Blanchard, 1989) and more recently in mice (Arakawa et al., 2007, other) where ethological aspects (e.g. day/night fluctuation of activity spatial distribution, place preference) and social and non-social behaviours of group-housed, freely-moving individuals can be continuously observed and quantified (Alonso et al., 2020, 2021).

In an effort to model in mice real world phenomena and pathological mechanisms as they could exist in the everyday-life of human patients (McCloskey 2011), we chose to apprehend the individuals’ behaviours and the group dynamic of Tph2-deficient and wildtype (Tph2-control) mice directly in their housing environment in a new version of the VBS which was designed and built for this project. The aim of this study was to characterize the impact of a complete lack of central serotonin in the everyday expression of social and non-social behaviours and their correlations in undisturbed, group-living mice of same-genotype, their spontaneous levels of activity and spatial distribution within the VBS over days. Training a machine learning algorithm (Random Forest classifier) on this high throughput dataset we identified key variables dissociating one genotype from the other when living in such ethologically-relevant environment. Because excessive aggression does not only drastically affect the life of the perpetrator but simultaneously affects the dynamic of the group and its structural organization, we evaluated if and how, compared to the control animals, the specific profile of the Tph2-deficient mice influenced the group’s structure and the dynamic of its hierarchical organization using SNA and the Glicko rating methods respectively.

**2. Materials and Methods**

***2.1 Animals.***

All mice were born at the animal facility of the Max Delbrück Center. The animals were housed in temperature (23±2°C) and humidity (50-70%) controlled rooms, lived under a 12h/12h non inverted dark/light cycle (light off at 18:00) and had *ad libitum* access to food and water throughout the project. Ten groups of four Tph2-deficient or wildtype male mice (*ntotal*=40) were used in this project: five groups of wildtype (Tph2-control) and five groups of Tph2-deficient mice on C57BL/6N genetic background (Alenina et al., 2009; Mosienko et al., 2012). Other C57BL/6N male mice (*n*=6) were used as the unfamiliar mice in the Three-chamber test.

***2.2 Experimental design.***

~~Due to the poor maternal care that Tph2-deficient mothers provide their pups, in order to increase the chance of survival of the Tph2-deficient pups co-housing of two mothers giving birth at the same time After weaning, littermates of same gender (but possibly different genotypes) were kept together in one large regular home-cage with filter top.~~ At six weeks of age, four mice of the same genotype were transferred to the same cage and were individually marked with unique Radio Frequency IDentification tags (RFID: 12 × 2.1 mm, 124 kHz, Sokymat, Germany, subcutaneous implantation in the scruff of the neck and under short (10s) isoflurane anesthesia). Afterwards the animals’ activity and health were regularly monitored (1h and 3h after marking and every day on the following days). At seven weeks of age, four mice from the same home-cage were transferred into in a new version of the VBS which was designed and built for this project for six consecutive days (or five days of 24h). Twenty-four to 48 hours after leaving the semi-automated VBS, the mice were tested in the Three-chamber test (one mouse after the other) during the light phase.

***2.3 Ethics statement.***

The experimental procedures described here were designed to allow for maximal animal welfare. Animals lived undisturbed as a group within their home-cage. Briefly, data collection was performed using automated observational methods applied to undisturbed group-housed animals. The health of the animals was monitored daily. Due to the observational nature of the study, the experimental procedure did not cause any damage, pain, or suffering to the animals. The animals were not sacrificed at the conclusion of the study. This study was performed under the supervision of the animal welfare officer (Tierschutzbeauftragter) heading the animal welfare committee at \_MDC\_ that approved the procedures.

Experiments followed national regulations in accordance with the European Communities Council Directive 10/63/EU.

***2.4 Semi-automated Visible Burrow System.***

***2.4.1 Material.***

The semi-automated VBS was designed and built following the description of the Visible Burrow System used in Arakawa et al., (2007), Figure 1. It consisted of a large regular home-cage (P2000, Tecniplast, Italy) where the different compartments were built in. The cage was separated in two (open and burrow) areas by a dark wall (PVC, 35 x 9 x 0.7 cm). The open area was a square (39 x 40 x 72 cm) delimited by the dark PVC wall and three transparent walls of the cage (22 cm high) on top of which extra white walls (particle wood board with smooth white finish, 50 cm high) were added. The extra white walls kept the area bright, prevented escapes and blocked most of the outside view of the cage. On top of the open area, was set a transparent Plexiglas lid which kept the illusion of openness of the area. The lid was tilted to avoid light reflections on the videos and had ventilations holes on the side (Plexiglas 44 x 38 x 14.5 cm). In the open area, regular food chow and water were available at two apertures (12 x 4cm for food) on opposite sides (Fig. 1). The bottom of the open area was covered with bedding (0.5 cm thickness). The other area, called the “burrow area” consisted of two separated dark chambers (PVC, 8 x 13 x 6.5 cm), connected to the open surface by transparent tunnels (Plexiglas). Chamber 1 had one straight tunnel (4 x 5 x 3 cm) connected to the open area while chamber 2 had two tunnels (straight: 4 x 5 x 3 cm and L shaped: 8+13 cm long x 5 x 3 cm) leading to the open area. A black plate covered the entire burrow area (burrows and tunnels; infrared (IR) transparent Acrylglas, 18 x 38 x 0.8 cm) and the three transparent walls of this side of the large home cage were tapped with black vinyl film so that the chambers and the tunnels were in near to complete darkness (Fig. 1). A grid of 24 RFID transponder readers (ID grid; Phenosys, Germany) placed under the VBS cage provided automated, continuous and simultaneous spatio-temporal information on each RFID tagged-animal present above. An infrared black and white video camera and two infrared lights were placed above the VBS cage. On the videos all animals were visible from all places in the VBS and in both light and dark phases. The ID grid and the video camera were connected to the same computer and data saved on an external hard drive for later manual analysis of the animal’s behaviours. The VBS cage could be easily disassembled/reassembled for cleaning of the parts in contact with animals.



**Figure 1. Illustrations of the semi-automated-VBS.** **(A)** A large rat cage was placed on top of a grid of 24 RFID readers and its walls were topped by extra high white walls. A camera was placed on top of the cage and aligned with the edge of the vertical partition (IR transparent, dark) of the burrow/open areas (drawn on SketchUp.com). **(B)** Schematic top view of the VBS cage with B1: one tunnel burrow, B2: two tunnel burrow, w: external water bottle with spout inside the cage, \*: food access zone with outside food chow reservoir.

***2.4.2 Method***

Each mouse was weighed before entering and after leaving the VBS cage. Each group of four mice spent five experimental days (five dark and five light phases) spanning over six astronomical days in the VBS cage. They entered the VBS at the onset (or a maximum of 30 min before) of the dark phase on experimental day 1 (at 18:00) and were removed from the system at the end of the light phase of experimental day 5 (after 06:00). In the VBS the mice were left undisturbed (e.g. no bedding change) and water and food were available *ad libitum*. The wellbeing of the animals was checked daily by inspection of the animals’ fur, posture and locomotion through the clear walls of the cage and by evaluation, on video, of their level of activity during the previous dark phase. After the VBS, mice were placed back together into the same (empty) regular home-cage they used to live in before.

***2.4.3 Data acquisition in the VBS (semi-automated)***

***2.4.3.1 Automated collection of RFID data, videos and identification of individuals.***

An RFID event was automatically recorded each time the RFID transponder of a given animal was detected by a RFID reader. For each event, the control-program (PhenoSoft Control program, PhenoSys GmbH) specified the date and time, duration of the event, the identity of the detected mouse and of the activated RFID reader. Events were continuously collected and saved (.csv) for the entire duration of the experiment. Thirty-second-long videos were recorded every 10 minutes during the five experimental days (CamUniversal software, Crazypixel). On each video, colored dots (one color per animal) were superimposed to the images of each mouse to allow visual identification of each individual of the group (Kolonikaefig software, PhenoSys). Marking the animals on the videos instead of color-marking their fur or ears was a more accurate, less invasive and also simpler method for long term identification of individuals within a group (Lewejohann 2010, Arakawa 2007).

***2.4.3.2 Manual annotation of behaviours (behavioural ethogram).***

Similar to former studies (Pobbe 2012), only the videos of the first four hours of each phase (dark and light) of all experimental days were analyzed (25 videos per phase, two phases per experimental day and five experimental days = 250 videos to analyze per group). Each time a mouse expressed one behaviour listed in Table 1, 1) the type and 2) duration of the behaviour, 3) where it took place in the cage and 4) the identity of any other mouse the focal mouse was interacting with during this behaviour (e.g. the mouse “m1” is “chasing” for “5 sec” in the “open area” the mouse “m4”) was reported in a behavioural ethogram (.xlsx). One focal animal was observed at a time, all four animals were observed per video. The videos were scored by two observers (MR and AH) trained to specifically and similarly recognize the behaviours described in Table 1. The same observer scored all videos of a given group of mice. During video-scoring, the observer was blind to the genotype of the group. Consistency between observers was evaluated as such: one observer would randomly select 10 to 20 videos of a group she did not yet annotate, score these videos and compare her results with the other observers’ results. If results differed, the two observers discussed discrepancies and adjusted their scorings’ strategies accordingly.

|  |  |  |
| --- | --- | --- |
| **Domain** | **Behaviour** | **Description of the behaviour of the focal mouse** |
| **Social behaviours (involves at least two individuals)** | | |
| Affiliative | **Allogrooming** | The mouse is licking or grooming another mouse |
| **Huddle** | The mouse is lying and/or sleeping in contact with another mouse (only scored in chambers) |
| Defensive | **Flight** | The mouse moves rapidly away from an approaching animal |
| Offensive aggression | **Chasing** | The mouse moves rapidly towards a fleeing animal (faster than following) |
| **Contact** | The mouse moves towards another animal from a distance. The movement ends with a physical contact with the target animal (not counted in chambers). |
| **Struggle** | The mouse moves towards another animal and rolling-and-biting follows. |
| **Struggle at feeder** | At feeder only. The mouse pushes another mouse without boxing or biting (otherwise counted as “struggling” or “biting”). Two separate “struggling at feeder” are counted when there is an in-between pause >1s. If pause is <1s, only 1 “struggling at feeder” is counted. |
| Mounting | The mouse mounts on the back of another mouse and shakes its hip quickly (similar to sexual behaviour). |
| Biting | The mouse is clearly biting another animal. |
| Social approach/ communication | **Approach to front** | The mouse moves towards the front of another animal from a distance. |
| **Approach to back** | The mouse moves towards the back of another animal from a distance. |
| **Following** | The mouse follows, at walking speed, the trajectory of another animal (slower than chasing). |
| Sniffing | The mouse is sniffing another animal. Sniffing is not counted in the chamber and when it is done while general exploration of the environment. |
| **Other behaviours (alone or in proximity of others)** | | |
| Maintenance | Drinking | At the water spout, the mouse is drinking (repetitive head and tongue movements directed toward the spout) |
| Eating | At the feeder, the mouse is eating (repetitive pulls with the head from the magazine or paws) |
| **Grooming** | The mouse is licking its own fur/paws/tail (repetitive head bobs) |
| Activity | Moving | The animal moves from one location to another but cannot be detected as another type of movement |
| Immobile | The animal shows minimal amplitude of movement to no movement |

**Table 1. Description of social and non-social mouse behaviours.**

**Notes Table 1:** Behaviours are grouped by domains (Lewejohann 2010). In bold are behaviours classically scored in VBS studies in mice (Arakawa 2007), in normal font are other behaviours often scored in social studies. The following behaviours: “being bitten”, “being sniffed”, “being mounted”, “being groomed” were scored but not analyzed to avoid redundancy with “sniffing” and “allogrooming”.

***2.4.4 Data analysis***

An experimental day consisted of 12h of dark and 12h of light phases starting at the onset of the dark phase. An experimental day spans over two astronomical days with the dark phase lasting from 18h of the first day to 05h59 of the next day.

***2.4.4.1 Activity in the VBS.***

The distance traveled per hour for each mouse was calculated from the event-based data generated by the ID-grid software (Phenosoft and Phenosoft Analytics, PhenoSys GmbH, Berlin). Distance traveled per hour is an indicator of the animal’s spontaneous activity over time. Due to a technical problem, data of four Tph2-control animals are missing from the dark phase of experimental day 4.

***2.4.4.2 Place preference in the VBS.***

Place preference in the VBS was calculated after the data generated by the ID-grid software (Phenosoft and Phenosoft Analytics, PhenoSys). The relative frequency (%) of activation of each RFID-reader of the 24-RFID-reader-grid located under the floor of the VBS indicated the relative preference of the animals (averaged per phase and per genotype over all experimental days) for each 24 zones of the VBS. In the open area of the VBS, four zones can be distinguished: a zone with access to the feeder, a zone with access to the water spout, a “safer” zone close to the separating wall and entries to the burrow area and a more “risky”, central zone (Fig.1 and Fig. 3). In the burrow area, half of the readers were located under one burrow and its two tunnels and the other half were located under the other burrow and its one tunnel (Fig. 3). A 24-tiles heatmap represents the spatial distribution of the 24-RFID reader. The darker the color of a tile, the more the corresponding RFID-reader was activated (relative to the activation of all the other readers of the grid (%)) by the presence of animals above it and thus the greater the animals (in average) preferred this location in the VBS.

***2.4.4.3 Social and non-social measures in the VBS***

For each behaviour described in Table 1, the total number of occurrences per animal, per genotype and per phase were analyzed (Fig.4). The effect of the VBS housing on the weight of the animals was calculated as the difference of weight before/after VBS. Per genotype, relationships between social and non-social behaviours were evaluated using a correlation table (Supplementary Tables 1-2).

***2.4.4.4. Random Forest Classification for differentiation of genotypes***

Machine learning was used to identify which variables, from all of the variables extracted during VBS housing (behaviours and activity, Fig. 4), were key to differentiate the Tph2-deficient from the Tph2-control mice. To this end we trained a Random Forest (RF) classifier (Breimann 2001) using the R package “randomForest” (Liaw 2002) and extracted the implemented feature importance (Gini index) for further analysis. We evaluated classification performance using a simple accuracy metric based on leave-one-out cross validation to ensure feasible features with regard to differentiability of genotypes. The input variables for the RF classifier were the total number of occurrences of each behavior illustrated in Figure 4 during the dark or light phase separately and the total distance traveled during VBS housing, for each animal of each genotype.

The estimated labels (i.e. genotype) of the test dataset were compared to the true labels of the animals and the overall accuracy of each classification, i.e. the ratio of correctly classified animals, was computed. The Gini index was assigned to each variable for its potential in differentiating the genotypes in each classification step. For robust results of the Gini index and RF we ran this procedure 100 times and reported the average (±SD) classification accuracy. A Gini index of 1 was the chosen threshold to separate behaviours highly indicating of a difference between genotypes from the other behaviours.

***2.4.4.5. Dynamic organization of the groups***

1. ***Evolution of aggressive and affiliative relationship strength between pairs of group living individuals using Social Network Analysis.***

For each Tph2-deficient and Tph2-control social network a node represented an animal (four nodes per network), a line between two nodes (a dyad) the occurrence of at least one interaction between them and the thickness of the line (the total number of interactions between a dyad) the strength (weight) of their relationship (weighted directed network). The higher the number of interactions was, the thicker the edge between the respective animal. For each selected behaviour the evolution of its social network was evaluated per day (dark phase only as more occurrences of behaviours happened during this phase).

This study explored the validity of different behaviours to be represented as networks before converging on “struggle at feeder” (prominent aggressive behaviour) and “allogrooming” (affiliative behaviour) as being the most informative type of interactions for the purpose of our network analysis. Following the hypothesis of structural differences between genotypes due to increased aggression in Tph2-deficient groups, we evaluated several groupings of distinct aggression-related behaviours including but not limited to “approach to back” AND “contact” AND “chasing” but found that the reduction to single behaviours delivered more robust results with regard to differential information). A range of standard network parameters were further explored such as clustering coefficient, betweenness centrality and shortest path length (Rubinov and Sporns 2010). The features of the current dataset however, in particular its small batch size of four animals per network, limited the information gain of including such parameters. The SNA of this study therefore focused on the dynamic (or daily) overall strength of both “struggle at feeder” and “allogrooming” networks in the VBS. This fundamental parameter reflects the role of a single animal or its “Relationship strength” within a network.

The overall interaction strength was assessed as a node’s (a single animal) total number of interactions within the directed network. We focused the analysis on overall strength instead of incorporating the directed versions of in-strength and out-strength due to the lack of variance between those parameters in the observed data.

1. ***Emergence and stability of hierarchy using Glicko-rating method and power distribution within groups.***

The individuals’ social rankings established by the Glicko rating system (Glickman 1999) has been found to highly correlate with other methods for dominance ranking (i.e. David’s scores and Inconsistencies and Strength of Inconsistencies (I&SI) ranking for instance in So et al., 2015). The clear advantage of the Glicko rating metric is to report on the dynamic changes in individual dominance ratings for each of the dyadic interactions within a group (So et al., 2015, Williamson 2016). Briefly, the Glicko analysis calculates individual ratings based on the outcome (i.e. win or loss) of each agonistic interaction between two animals. If an animal won a given interaction its rating increased while the rating for the losing animal decreased. The Glicko rating model (PlayerRating R-Package), as an extension of the Elo dynamic paired comparison model as presented in Neumann et al., (2011), did not only iteratively compute the animals’ rank but also the standard deviation of its ranking history to get an estimation of the certainty of the rating, which is again used to update the animals ranking. Additionally, this ranking model did not only update the given rankings when a direct interaction occurred but also when other dyadic interactions occurred, recognizing the group as a network more than a sum of separated pairs of individuals. Following Williamson et al., (2016) we set the initial ranking and certainty values equal for all animals. Differing from Williamson et al., (2016) we set the ranking and certainty equal to zero to enable negative ratings to improve the visualization of social hierarchy. We further set the ranking update constant equal to 1, which has been shown to create little impact on the final results (So et al., 2015) and still represent a sound value for mouse agonistic interactions (So et al., 2015). The Glicko rating analysis and power distribution analysis were performed on data “struggle at feeder” since this subset of data has shown more robust rank order across time than grouping of several aggressive behaviours (So et al., 2015).

We used the Glicko rating system (Glickman 1999) to appraise 1) if group stratification or hierarchy [i.e. leveling of final ranking of individuals above and below the initial rank with the dominant animals having higher rank] could be observed in VBS housing, within groups of Tph2-deficient mice and groups of Tph2-control mice, 2) if one distinct dominant animal emerged at the end of the test, 3) how individual hierarchical ratings dynamically evolved over time, and 4) how rapidly and stable dominant males emerge within each groups [i.e. Glicko Rank history correlation]. The stability of the dominance hierarchy within a group (i.e. when the final hierarchical ranking emerges in time) was evaluated by computing spearman’s rank correlation of the final glicko rating with intermediate ratings during update phases in the glicko history. For comparability reason the rating history was divided into 20 intervals representing 5% of the overall updates of the glicko ratings. High correlation values in early intervals represent a good approximation of the underlying social hierarchy in time, either due to emergence of a distinct dominant animal (e.g. Batch 7) or due to a continuous power struggle within the batch (e.g. batches 9 and 10). Finally, we evaluated how inequitable the distribution of power could be within same-genotype groups of mice, with and without central serotonin. Here we evaluated the relative proportion of power the alpha male is imposing on the beta male compared to the most subordinate animal. This can be represented as the ratio of difference in rating between alpha and beta male over the difference in rating between alpha and most subordinate male. A high value represents a strongly despotic alpha male imposing relatively similar amounts of power towards all other animals.

These analyses were performed on the output of the video data described above. This data is not continuously collected due to breaks in video recording. This may introduce minor inaccuracies in the history of dyadic interactions but should be evened out by the inclusion of rank certainty in the glicko rating algorithm.

***2.4.4.7 Criterions for pathological aggression***

Based on Takahashi et al., (2011), for each individual, we used from the behavioural ethogram the three following quantitative parameters: (1) the latency to first attack (filtering for “struggling” or “struggling at feeder” or “chasing”), (2) the frequency and (3) the mean duration of attacks [Miczek et al. 2002, 2003 in Takahashi 2011]. A short latency to attack associated with increased frequency and duration of attacks would suggest escalated aggression in mice (Table 2). Other qualitative aspects of abnormal aggression were more difficult to extract from our data. The body location of bites (especially to vulnerable body parts) could not be assessed in our study as very few biting behaviours were observed or skin wounds were found. The lack of ritualistic behaviours (Haller 2005) could only be indirectly and tentatively measured as the (4) ratio of fight/threat behaviours (Fight: struggle + struggle at feeder and Threats: chasing + following + atb). The theory be that the smaller the ratio, the more the threats were able to stop aggression. The lack of responses to appeasing signals could not be evaluated from the angle (top view) of our videos. Finally and although the conditions in the VBS did not allow to evaluate if attacks were context independent, such as “aimed at the opponent regardless of its sex or state (free-living/anaesthetized/dead) or the environment (home/neutral cage)” [Koolhaas 1978], we could spatially locate where aggressions happened the most and if these places were appropriate places for such behaviour (Haller 2005; Table 2).

***2.5. Three-chamber test***

***2.5.1 Material.***

The apparatus consisted of a rectangular white box (60 x 40 x 22 cm) divided into three equal-sized chambers (20 x 40 x 22 cm). Dividing walls were made from clear Plexiglas, with rectangular openings (9 x 0.8 x 12 cm) allowing free access to each chamber. Two clear Plexiglas doors were used to block the openings when needed. Two round metal-wired cages with grey PVC covers at the bottom and top (Ø10 cm x 21 cm) were used. In each metal-wired cage could be placed one stranger mouse. Once the metal-wired cage placed in the center of a lateral chamber, the distance between the wires allowed the subject and the stranger mice to see, hear, smell, and touch each other but prevented fighting. A video camera fixed above the apparatus allowed to record the position and behaviours shown by the subject mouse at any time and in the entire apparatus. All videos were saved on a computer for later analysis.

***2.5.2 Method.***

The subject mouse was weighed before entering the Three-chamber apparatus. The stranger mice were kept in a separate experimental room, and only transferred to the testing room when needed. The stranger mice had been previously habituated to the metal-wired grid (10 minutes daily, for at least three consecutive days before the testing day). The Three-chamber test consisted in three phases (Habituation, Social Preference and Social Recognition). In the Habituation phase, the test subject was first placed in the middle of the chamber and allowed to explore this chamber for five minutes (the access to the lateral chambers were blocked by transparent doors). Then, the lateral doors were removed and the mouse was allowed to explore the entire apparatus for 10 more minutes. At the end of this Habituation phase, the mouse was gently pushed back into the center of the apparatus and accesses to the lateral chambers were blocked. In the Social Preference phase an empty metal-wired grid was placed in one lateral chamber and a metal-wired grid with an unfamiliar mouse that had no prior contact with the subject mouse (stranger 1) in the other lateral chamber. The location of the unfamiliar mouse in the left vs. right side chamber was systematically alternated between test animals. After the lateral doors were removed, the subject mouse could explore the entire apparatus for 10 minutes (Social Preference test). The subject mouse was then gently guided back into the center of the apparatus, accesses to the lateral chambers blocked and the two metal-wired grids removed from the apparatus. After an inter-test interval of 5 minutes, the lateral doors were opened and the subject mouse was allowed to explore the entire apparatus for 10 more minutes (Social Recognition test). During the Social Recognition test an unfamiliar mouse (stranger 2) was placed in the previously empty metal-wired grid. The grid with stranger 1 was placed back into the same lateral chamber as before. The mouse had a choice between the first, already-investigated mouse (now-familiar mouse) and the novel unfamiliar mouse. At the end of the Social Recognition test, the subject mouse and the metal-wired grids were removed. The three chambers and doors were cleaned with ethanol (70%) and the metal-wired grids (emptied from the stranger mice) wiped cleaned with water and dried.

***2.5.3 Data acquisition and analysis.***

Videos of the tests were recorded and saved for offline analysis by the video-tracking system Viewer 3 (Viewer, Biobserve). In the Three-chamber test, both Social Preference and Social Recognition, were measured as the total time spent in each chamber (%) and in close proximity with the grid-cage per 5 min bins.

***2.6. Statistical analysis***

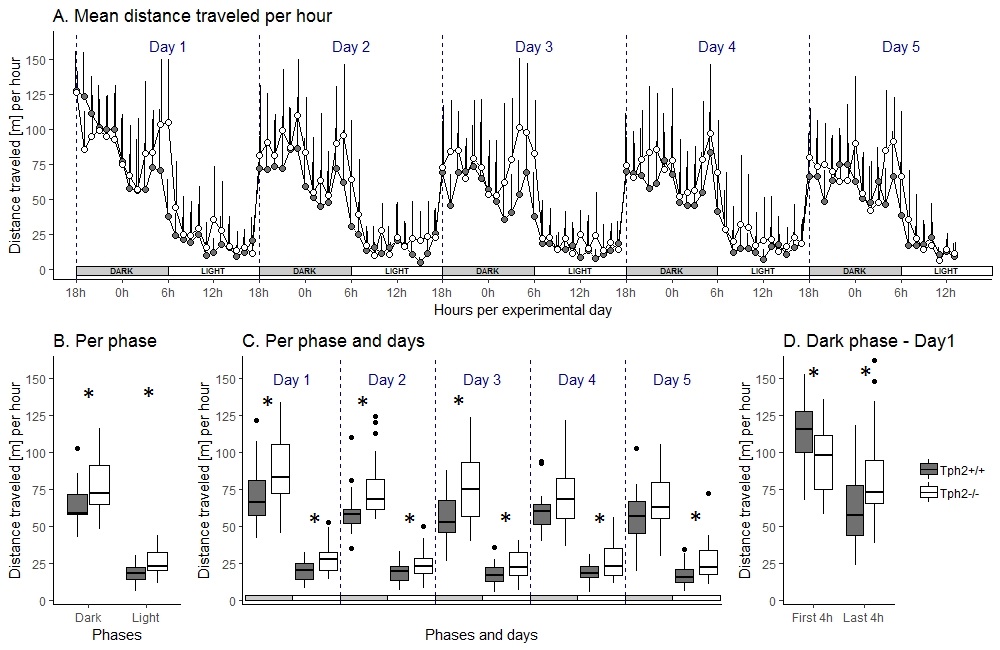
We performed two types of statistical analysis. We validated the continuously-collected-RFID-data with regards to influence of experimental (genotype) and random (animals, batch) factors using markov chain monte carlo simulations of general linear mixed models (MCMCglmm R-package, Hadfield 2010) and checked for a significant difference of the posterior distribution of the simulations with zero to assess the influence of the random variables. Other statistical tests were genotype based comparisons on a variety of experimental (distance traveled overall, per phase, per day; place preference; total behavior occurrences; parameters of pathological aggression and weight) and analytical (social network parameters) variables. To this end we computed the exact Wilcoxon-Mann-Whitney test with multiple comparison correction. The code base is written in the R programming language and made available via GitHub (https://www.github.com/patrikbey).

**3. Results**

**3. 1. Behavioural profile of serotonin-deficient mice in the natural-like environment of their VBS home-cage across days.**

***3. 1. 1. Activity and Place preference in the semi-automated VBS.***

Animals of both genotypes showed a similar pattern of activity across days. Their activity level drastically rose and fell at the onset of each dark (high-activity) and light (low-activity) phases across experimental days (Fig. 2A). Although the activity level was higher during dark phases than during light phases, the animals were also more active during the first and the last four hours of each dark phase with a twofold decrease of activity during the four hours in between. Their activity levels were low and constant throughout light phases (Fig. 2A). Despite this similar pattern of activity, Tph2-deficient mice covered longer overall distances than Tph2-control animals (MCMCglmm random factors as animal and group: pMCMC=0.011, post.mean=9.865, [l-95% CI = 2.375, u-95% CI= 16.857]) and in both phases (Fig. 2B., Exact Wilcoxon-Mann-Whitney test: all dark phases, Z = -2.4075, p-value = 0.01548; all light phases Z = -2.5427, p-value = 0.01031). More specifically, the Tph2-deficient mice were found consistently more active in all, except the last two, dark phases in the VBS (Fig. 2C., day1 Z = -2.3263, p-value = 0.01954; day2 Z = -2.9485, p-value = 0.002643; day3 Z = -2.4616, p-value = 0.01319). Tph2-deficient mice were also more active than the Tph2-control mice on most light phases (Fig. 2C., day1 Z = -2.2181, p-value = 0.02633, day3 Z = -2.0558, p-value = 0.04018, day5 Z = -2.1099, p-value = 0.03501). Surprisingly, after the animals entered the VBS for the first time, the Tph2-control mice appeared to be more active than the Tph2-deficient mice (first four hours of the dark phase) (Fig. 2D., Z = 2.3263, p-value = 0.01954), which was not the case during the later peak of activity of the same dark phase.

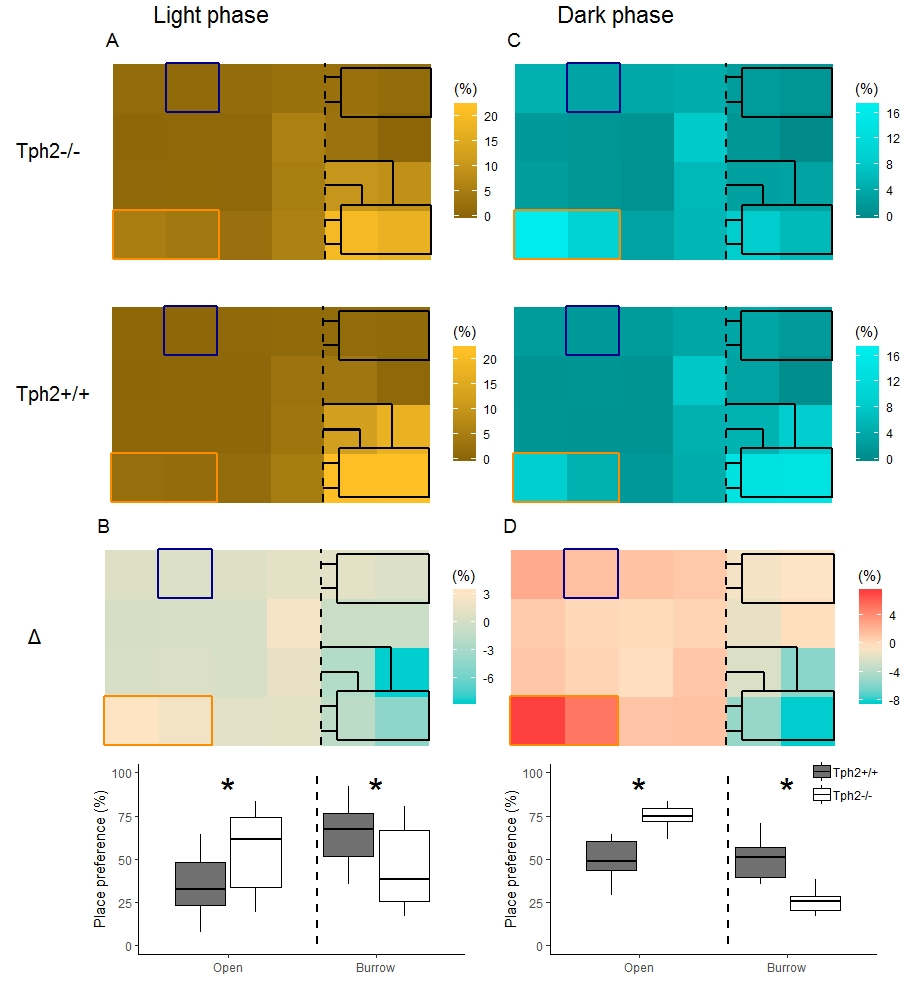
****

**Figure 2. Distance traveled in the automated-VBS.**

**(A)** Mean distance traveled per hour and per genotype (Tph2+/+: control mice, grey symbols, KO: Tph2-/-: deficient mice, white symbols), across days (+SD). Each experimental day starts (dashed line) at the onset of the dark phase (18h-5h59; grey box) and finishes with the end of the following light phase (6h-17h59; white box). **(B)** Boxplot of individual mean distance traveled per hour and per genotype, per phase. **(C)** Mean distance traveled per hour and per genotype, across days. **(D)** Mean distance traveled per hour during the first and last four hours of the dark phase in experimental day 1. Boxplots show median, quartiles, 5th/95th percentiles and outlying points (geom\_boxplot {ggplot2}, R). Exact Wilcoxon-Mann-Whitney test, \* p < 0.05. m, meter.

The heat-maps in Figure 3 revealed distinct diurnal and nocturnal spatial preferences for different zones within the VBS and between genotypes. During the light phase, animals of both genotypes were mostly detected in the burrow area with a clear preference for the one burrow with two tunnels (top half of the burrow area; Fig. 3A). Although the pattern of occupation of the different zones of the VBS seemed equivalent between the genotypes (similar locations with similar shades of colors per genotype), the difference between both heat-maps (∆ = [Tph2-/-] – [Tph2+/+]) indicated that Tph2-deficient mice spent less time in the two-tunnels-burrow and more time in the open zone than the Tph2-control mice during the sleeping (light) phase (Fig. 3B: Open area, Z = -3.8298, p-value = 8.898e-05). During the dark phase, mice of both genotypes were mostly detected at the feeder and close to the separating wall on the open side of the cage and more often in the burrow with two tunnels than in the burrow with a single tunnel (Fig. 3C). During this active (dark) phase of the day, the difference in occupation of these zones between the genotype was even more pronounced than during the light phase (stronger variations of colors on the ∆ heatmap), with the Tph2-deficient mice significantly more often detected in the open and especially at the feeder than the Tph2-control mice and less often in the burrow with the two tunnels (Fig. 3D: Open area, Z = -5.3018, p-value = 1.741e-10).

\*



**Figure 3. Place preference in the VBS during light (A-B) and dark (C-D) phases.** **(A)** Averaged place preference (%) of Tph2-deficient (-/-; top) and Tph2-control (+/+; bottom) animals during the light phase (all days). Each tile of the heat-map represents the position of one RFID-reader as it was located under the VBS (24-RFID-grid). Food (orange rectangle) and water (blue square) were available from two distinct zones in the open area. The vertical black dashed-line indicates the separation between the open and the burrow areas. The brighter the color of a tile is, the more the RFID-reader was activated (relative to the activation of all the other readers of the grid (%)) and thus the greater the animals preferred in average this location in the VBS. **(B-top)** Differences in place preference (%) between Tph2-deficient and Tph2-control (subtraction of heat-maps in (A)) during the light phase. **(B-bottom)** boxplot representation of averaged place preference (%) in **∆,** by zone (open *vs.* burrow). **(C)** Averaged place preference (%) of Tph2-deficient (top) and Tph2-control (bottom) animals during the dark phase (all days). **(D-top)** Differences in place preference (%) between Tph2-deficient and Tph2-control (subtraction of heat-maps in (C)) and **(D-bottom)** boxplot representation of averaged place preference (%) in **∆,** by zone (open *vs.* burrow). Boxplots show median, quartiles, 5th/95th percentiles and outlying points (geom\_boxplot {ggplot2}, R). Exact Wilcoxon-Mann-Whitney test, \* p < 0.05.

***3.1.2. Social and non-social behaviours in the home-cage***

In the semi-automated VBS, all behaviours listed in Table 1 were seen in both genotypes. Only “biting” (occurred five times in total (three times in Tph2-conrol and two times in Tph2-deficient mice)) and “mounting” (occurred six times (three times in each genotype)) were very rarely seen. All behaviours, except for “huddle” which is happening while sleeping, were most expressed during the active (dark) phase of the day (see scales in Fig. 4). Tph2-deficient mice performed significantly more offensive aggression such as “approach to back”, “chasing”, “contact”, “struggle” and “struggle at feeder” than the Tph2-control mice and in both phases (except for “approach to back” and “chasing” during the light phase, for which the occurrences of behaviours are too rare to be meaningfully tested; Fig. 4A: Dark phase, approach to back Z = -2.7593, p-value = 0.0048, chasing Z = -2.3208, p-value = 0.0196; contact: Z = -2.3702, p-value = 0.0169; struggle Z = -4.3708, p-value = 2.077e-06; Supplementary Fig. 1 Light phase, contact: Z = -2.0371, p-value = 0.0414; struggle Z = -4.3624, p-value = 2.657e-06). In Tph2-control mice, “chasing” and “approach to back” (both phases) and “struggle” and “struggle at feeder” (in light phase only) were rare behaviours (Fig. 4A, Supp Fig.1). Tph2-deficient mice were more defensive (“flight”) than the Tph2-control mice during the light phase (Supp Fig.1, Z = -2.1969, p-value = 0.0304) but not during the dark phase (Fig. 4B). Regarding social (non-aggressive) approaches, during the dark phase, Tph2-deficient mice performed significantly less “sniffing” behaviours (Z = 2.4759, p-value = 0.0124) than Tph2-control mice but did not differ in total number of “approach to front” or “following” behaviour during this same phase (Fig. 4C). These behaviours were rarely observed during the light phase in both genotypes and thus, were not compared statistically (Supp Fig.1). In both phases, Tph2-deficient mice were found eating and drinking (although “drinking” was more rarely observed, probably because of the shortness of the behaviour) significantly more often than the Tph2-control mice (Fig. 4D: Dark phase, drinking Z = -3.1292, p-value = 0.0012; eating Z = -3.5444, p-value = 0.0002; Supp Fig.1: Light phase, drinking Z = -3.1837, p-value = 0.0013; eating Z = -4.2074, p-value = 6.191e-06) and grooming during the dark phase was less often witnessed in Tph2-deficient mice than in Tph2-control mice (Fig. 4D: Dark phase, Z = 2.5875, p-value = 0.0087; Supp Fig.1: Light phase n.s.). Finally considering affiliative behaviours, Tph2-deficient mice showed less “allogrooming” behaviour than Tph2-control mice, in both phases (Dark phase, Z = 4.3102, p-value = 3.562e-06; Light phase, Z = 2.6884, p-value = 0.0064) while “huddled” as many times as the Tph2-control mice in both phases (Fig. 4E, Supp Fig.1).

The relationships between social and non-social behaviours were then explored per genotype (correlation tables in Supplementary Tables 1-2). With some exceptions, Tph2-deficient mice presented equivalent relationships between variables than Tph2-control mice. In both groups occurrence of “eating” positively correlates with the occurrence of “struggle at feeder” (Supplementary Table 1 and 2) but while “struggle at feeder” (or “eating”) were negatively correlated with “allogrooming” and “grooming” in Tph2-control group, “struggle at feeder” (or “eating”) and “contact”, were both negatively associated with the main affiliative behaviour of “huddling” in Tph2-deficient mice. Only in the Tph2-control group “huddle” negatively correlated with “chasing” and “struggle” (and “sniffing”). In Tph2-control groups, the occurrence of a “contact” was positively associated with the occurrence of an “approach to back” (atb) and to “chasing” and “struggle” (away from the feeder) while surprisingly, in Tph2-deficient mice, “contact” was negatively related to “atb” (and “huddle”), but positively related to “struggle at feeder” (and “eating”) (Supplementary Table 2). Only in the control group “chasing”, “contact” and “struggle” covariated positively. In the mutant group “atb” is the behaviour which correlated with most of the other scored behaviours. It was positively related to “huddling” and “allogrooming” but negatively to the occurrence of “struggle at feeder”, “grooming” and “eating”. In the Tph2-control mice “sniffing” positively correlated with “allogrooming” and “atb” and negatively with “huddle”. Finally, in the mutant group, “allogrooming” negatively correlates with “eating”.

Finally, Tph2-deficient mice lost more weight after five days in the VBS than the Tph2-control mice (Exact Wilcoxon-Mann-Whitney test: Z = -2.6436, p-value = 0.0072; Tph2-deficient weight: [before, mean: 18.8 ± SD: 2; after, 18.9 ± 3.4]; Tph2-control weight: [before, 19.8 ± 5.1; after, 20.4 ± 4.7]).

* + 1. ***Behavioural differentiation of genotypes using a Random Forest Classification.***

The training of the random forest (RF) classifier, on the social and non-social behaviours and distance traveled in the VBS of the Tph2-deficient and control mice, led to high precision in genotype prediction with an averaged accuracy of 81.4% (±1.2) over 100 runs. For each run of the classifier (number of animals \* 100) we obtained the Gini index for each input variable. These values provided a robust estimate for the importance of each of the given behaviours to differentiate Tph2-deficient and control mice (Fig. 4F). With the arbitrary chosen Gini index ≥1, the behaviours with the greatest potential for differentiation of the two genotypes were “allogrooming”, “struggling at feeder” and “eating” in the dark phase and “eating” and “struggle” during the light phase (Fig. 4F).



**Figure 4. Total number of occurrences of social and non-social behaviours in the home-cage. (A-E)** Total number of behaviours (grouped by domains) per genotype and over all experimental days during dark (upper graph of each panel) phases. ATB: approach to back; ATF: approach to front. **(F)** From the RF classifier, plots of one hundred Gini values, for each VBS variable, during the dark and light phases separately. Gini index >1 (dotted line) indicates behaviours highly different between genotypes. SAF: struggle at feeder. **(G)** Social networks for “struggle at feeder” (top) and “allogrooming” (bottom) of an exemplary group of Tph2-deficient (-/-) and of Tph2-control (+/+) mice across dark phases of successive days. A node represents an individual and the width of a line the overall strength of that behaviour between the pair of animals for the given day. Boxplots show median, quartiles, 5th/95th percentiles and outlying points (geom\_boxplot {ggplot2}, R). Exact Wilcoxon-Mann-Whitney test, \* p < 0.05.

* 1. **Role of serotonin in the dynamic organization of groups of mice in their home-cage and across days.**
     1. ***Evolution of aggressive and affiliative relationship strength between pairs of group-living individuals using Social Network Analysis***

By plotting the network of interactions happening during dark phases (when there are more occurrences of all behaviours) of two of the most discriminative variables between genotypes (one offensive: “struggle at feeder” and one affiliative:“allogrooming” type of behaviour) and in the form of a day by day directed graph, we could visually witness the dynamic of the quality of interactions between pairs of animals within their social network. As expected, the created networks demonstrated clear differences in topology between the Tph2-deficient mice and controls (overall strength across days, allogrooming, Z=-4.6798, p-value=1.4358e-6; struggling at feeder, Z =-4.2158, p-value =1.2446e-5). On a daily basis, Tph2-deficient mice compared to Tph2-control mice struggled at feeder with a higher interaction strength, from day 1 to day 5, (day1:Z =-3.6607, p-value =0.0001; day2:Z =-3.7961, p-value =7.3489e-5; day3:Z =-2.8738, p-value =0.0020; day4:Z =-5.1147, p-value =1.5709e-7; day5:Z =-4.277, p-value =9.4703e-6; Fig. 4G showing two representative groups of mice) but performed “allogrooming” with a lower interaction strength from day 2 to 5 (day2:Z =-5.1973, p-value =1.0111e-7; day3:Z =-3.3003, p-value =0.0005; day4:Z =-2.3549, p-value =0.0093; day5:Z =-3.8115, p-value =6.907e-5; Fig. 4G showing two representative groups of mice).

* + 1. ***Emergence and stability of hierarchical ranking using Glicko-rating method and power distribution within groups***

The changes in Glicko ratings over time of each individual of each group of Tph2-deficient and Tph2-control mice are plotted in Figure 5.A. The Glicko rating is updated after each observed agonistic interaction (struggle at feeder only) and based on the outcome (either win or loss, here the direction of the interaction defines the winning animal (initiator) and losing animal (receiver)) of the interaction. Because the total number of agonistic interactions varied between groups, the total number of interactions (displayed on x axis) is not the same between plots.

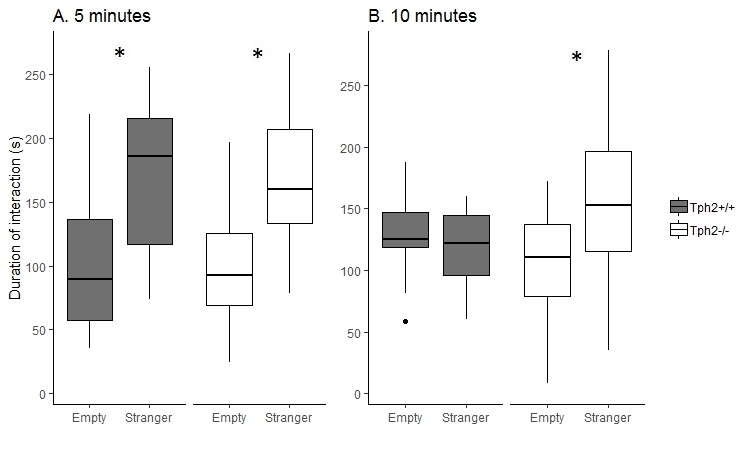
In all groups, one (in two groups) or two animals’ rankings (in the eight other groups) were found above their initial Glicko rating (y=0). These results indicated in both genotypes the existence at the end of the VBS of a stratification of the subjects’ ranks, into higher and lower ranked individuals (Fig 5.A). In all groups, except two (one of each genotype), at the end of the VBS one dominant individual clearly departed from the other individuals with the highest rating compared to them. In group 4 (Tph2-control) and group 10 (Tph2-deficient) one dominant animal could not be clearly identified, due to similar ranking dynamics of at least two individuals. In the Tph2-control groups the higher ranked individual (highest final Glicko rating at the end of a VBS) already emerged as the most dominant animal after a median number of 50 agonistic interactions (mean: 63 ± SD: 43; (group scores: 50, 20, 50, 100, 125)). The stability point is defined as where the line hits a correlation value of 80% for the last time, in case of bumps. In the Tph2-deficient groups, the dominant males at the end of a VBS emerged as the dominant individual of its group after a median number of 225 agonistic interactions (mean: 219 ± SD: 149; (group scores:130+25+320+400)). Taken together, these results suggest that dominance emerged more readily (after a smaller number of interactions) in the Tph2-control groups compared to the Tph2-deficient groups.

Finally, we examined the power distribution of the dominant males across groups by evaluating how each alpha male monopolized agonistic interactions within their social group or said otherwise, how (un)equally distributed was the power within groups? Considering the individual’s final rankings, a dominant male is considered “despotic” when it imposes towards the beta (second highest ranked) male a power of one third of the total imposed power on the lowest ranked animal. Here we found that 40% of the Tph2-deficient alpha males imposed such “despotic power” while 80% of the Tph2-control dominant males did (See group’s scores in Supplementary Table 3). The minimal power a dominant mouse imposed on the group was three times higher in the Tph2-control groups than in the Tph2-deficient groups (23.9% *vs.* 7.9%). At the same time the most despotic alpha males in both groups displayed a similar ratio of imposed power (around 67%; see group 1 and group 6 in Fig.5). Similarly, the average imposed power in the control groups surpasses the deficient groups by a factor of 1.5 (42% vs 30%). The observed variance in imposed power within the groups is twice as large in the Tph2-deficient cohort (0.058 vs. 0.029) indicating a less robust power distribution.

****

**Fig.5. Temporal dynamics of individual Glicko ratings for Tph2-control groups (labeled 1-5) and Tph2-deficient groups (6-10). (A)** The change in individual Glicko ratings over time. Each line represents the ratings of one individual while the solid black line represents the final alpha male. Ratings are recalculated for every individual after each agonistic interaction and are plotted on the y axis against the ranking updates on the x axis. **(B) Rank history correlation (%).** At each update step the correlation with the final ranking is computed. Each time interval representing 5% of the overall interactions shown cumulates the present correlations in the boxplots. High correlations in early intervals indicate the emergence of clear social hierarchies (e.g. 7). Medium correlation values indicate continuous struggle to establish clear hierarchy (e.g. 5).

* 1. **Social cognition in the three chamber test**

****

**Figure 6. Three chamber test and social preference. (A)** The first 5 minutes of test. **(B)** The last 5 minutes of test.

In the first five minutes of social preference test both groups of mice prefer interacting with the unfamiliar individual (Tph2-deficient mice: W = 24, p-value = 0.0006612; Tph2-control mice: W = 16, p-value = 0.03147). In the next five minutes while Tph2-control mice lose interest for the now “familiar” animal, the Tph2-deficient mice keep interacting more with the mouse than with the empty cage (Tph2-deficient mice: W = 42, p-value = 0.01675). Tph2-deficient mice do not habituate as fast as the Tph2-control mice to a new individual. They keep investigating the novel individual over the 10 minutes of test while the controls cease to prefer the novel individual after 5 minutes of test.

* 1. **Pathological aggression in mice**

The first attack in Tph2-deficient mice occurred sooner than for the Tph2-control mice. They attacked (struggle and struggle at feeder together) with a higher frequency and used significantly less warning signals (e.g. threats: chasing, following, atb) than the control mice. However, the duration of a given attack was not longer for the Tph2-deficient mice and fights occurred mainly at the feeder (Tab. 2).

**Table 2. Pathological aggression in mice based on Takahashi 2011 and Haller 2005.**



**Notes.** KO: Tph2-deficient mice, WT: Tph2-control mice, med: median, m±sd: mean ± standard deviation, n/a: not applicable, n.s: not significant, p value after Mann-Whitney test.

**Discussion**

In this study we performed an in-depth analysis of home-cage behavior of group living mice, in order to broaden our understanding of the role of central serotonin (or its lack thereof) in the expression of everyday-life aggression, social and non-social activities and of group’s dynamic organization (expression of daily dynamic interactive complex interactions). In this home cage context, although Tph2-deficient mice presented some well conserved mouse characteristics they too presented important behavioral anomalies, altered social network characteristics and different dynamics of group formation.

In the undisturbed ethological-like conditions of the VBS housing, Tph2-deficient mice showed typical day-night fluctuation of activity and equivalent utilitarian use of the different zones of the cage depending on the phase of the day. Similar to controls, they visited more the food and open area during active phases and the safe and sheltered areas during sleeping phases. The congenital lack of serotonin in these mice did not appear to impact the circadian-associated phenomenon of phase-shift of activity in response to environmental light changes. Despite known interactions between serotonin and circadian systems (regulation sleep-wake cycle) and their respective roles in the expression of seasonal mood disorders for instance (refs), this result confirms the role of a larger neurobiological network for the regulation of these processes. The lack of congenital serotonin did not abolish the expression of any specific behavior, but drastically affected their occurrences, except for the life-essential affiliative behavior of “huddling” (i.e. “sleeping in direct contact with at least one other mouse”) that Tph2-deficient mice performed as often as control mice. The expression of this important behavior for maintenance of group cohesion (ref) was well preserved and was also negatively correlated with expression of offensive behaviors in both groups. Finally and despite prior uncertainties on the ability of Tph2-deficient violent groups to organize hierarchically, classical indexes of hierarchy were found comparable between genotypes. The conservation of typical daily life characteristics in animal models is as important as observed deficits for it gives a more comprehensive face validity to the model. In their everyday life, violent patients are not socially maladapted in all contexts and all endeavors, they can also blend in some layers of the society.

Still, life without brain serotonin and in interaction with other “similar” mice had a deleterious impact on most other individual and group characteristics. Pathological aggression has been described as inappropriate, frequent and prolonged bouts of aggression which lead to increased burden for the animals (Haller 2005). On the five criterions we could analyze in our study, Tph2-deficient mice presented only three of them which does not fully qualify these mice as a model of pathological aggression as it is defined by these criterions (Tab. 2). They attacked sooner, more often and displayed fewer “warning” signs than controls but a fight would be of similar duration and rarely occurred in inappropriate zones of the cage (e.g. no fights in burrows). Indeed, aggression in all forms (ATB, chasing, contact, struggle and struggle at feeder) and defensive (i.e. flight) behaviors were highly displayed by Tph2-deficient mice and especially at the feeder. Not only did Tph2-deficient mice fought more (at the feeder) individually but each mouse would struggle at feeder with all the other mice of the group (ties between all pairs), repeatedly (ties are thick) and consistently over days. In contrast to the Tph2-control networks where struggling at feeder was observed between fewer and varying pairs of mice over days.

Food as a resource is a natural trigger of aggression. If aggression at feeder is not abnormal, the absence of aggression de-escalation with time is. Tph2-deficient mice are known to have a higher metabolic rate (ref) and during VBS housing although they were found eating and drinking more often, they did not gain as much weight as control mice. Serotonin through a diversity of serotonin receptors plays an important role in the inhibition of feeding behavior via the reinforcement of satiety signals (ref PDD) and in the control of impulsive behaviors targeting food (ref PDD). Considering their high level of activity and fights which are energetically expensive, these mice might have been particularly motivated to feed regularly although it increased their chances to meet and potentially fight with other mice at feeder. The presence of another conspecific at feeder can be interpreted as a provocation prone to elicit aggression (Blanchard 2003-stim elicit aggression=response to provocation). On the long term however, persistent fights at feeder and between all mice are not seen in control mice as they flexibly adjust their behavior with time to limit the burden of such fights (see network analysis). Lee et al., (2018) even observed individual mice modify their feeding times according to the feeding times of other mice with which they had history of aggression. In rodents and primates, subjects with lower levels of 5-HT are more likely to persevere in actions no longer favorable (2 ref PDD) and serotonin and its multiple receptors are essential players in the control of behavior, perseverative behavior, behavioural flexibility and the extinction of context-dependent conditioned behaviors (PDD, Dellu-H 2018, Robins; behavioural flex: (Lewejohann 2009 or 10 5-HTT KO mice in vivarium). Interestingly, although Tph2-deficient mice were unable to withhold aggression in the specific context of the feeder, in the context of the burrow they could restrain themselves from being aggressive and tolerate each other presence to huddle for certain periods of time. In their undisturbed home cage environment, aggression in Tph2-deficient mice appeared to be highly context-dependent (i.e. food dependent) and their inability to de-escalate, aggression revealed poor behavioral inhibition and flexibility in this context specifically.

In the current study, mice of same-genotype were grouped together in order to control for any beneficial or deleterious effect control mice could have had on the behavior of the mutant mice (refs). Such reduced diversity of phenotype within human groups is rare and in the specific contexts of prison, gangs or specific jobs (trading, professional sport…) high aggression for power is often an issue. In genetically “identical” mice it has been shown that with time and through social interactions individuality always emerged and that these individual differences were also sustained by long lasting biological changes (Freund, Faure mouse city, ppr twitter/phone). In line with these observations, the very aggressive “identical” mice of the current study could still dynamically organize their groups into individually stratified and stable hierarchies (dynamic and final Glicko ratings) although hierarchies emerged later, and the power of the alpha male was less despotic and more diffused than in control groups. These results highlighted a non-essential role of serotonin to build up a social hierarchy but its specific role in structuring and dynamic aspects of the group formation (SNA, diffused power, late emergence of leader).

The emergence of a dominant individual is a dynamic process which rests on the level of social competence of the animals of the groups. Using different sensory systems, the subordinate individuals monitor the behavior of the alpha male and with such social knowledge can adjust their behaviours and display appropriate protective social behaviours, inhibit aggression and avoid conflict (Curley 2016). To show such social competence mice must not lack sensitivity to social cues, be delayed in encoding and forming social memories or be unable to flexibly adjust behaviours. Knowing that serotonin boosts the reinforcing value of social signal (Kalenscher), following the “for better for worse model” (Kiser 2012), a lack of serotonin in the Tph2-deficient mice could dampen their sensitivity to social and non-social cues present in their environment.

After entering the unfamiliar environment of the VBS, Tph2-deficient mice showed a transient although prolonged suppression of exploratory behavior. This unusual behavior was not due to social interferences or affective reasons. Mice entering the same VBS were familiar with each other and did not fight more during the first four hours in the VBS (data not shown) than later. TPh2-deficient mice did not present anxious-like tendencies as they spent relatively more time at the feeder and thus in the risky (open) area of the VBS than controls throughout VBS housing.

It is very unusual for a mouse to not explore intensively an “unfamiliar” home cage and similarly odd to keep investigating an “unfamiliar” mouse (in the three chamber test) for twice as long as a control mouse would. Serotonin is an emergent target to counter memory alteration through lack of synaptic plasticity (ref). Serotonin is required for memory formation (and especially short and working memory- see Hritcu 2007; Gonzalez-Burgos 2008) and encoding of familiarity and phenomenon of déjà vu (Kalra 2007). Tph2-deficient mice might be impaired in detecting the unfamiliarity of their new home cage or of a conspecific and in forming new memories necessary for rapid adjustment of their behaviours.

Along with aggression and feeding behavior, the Random Forest classifier ranked “allogrooming” and “sniffing” among the four most discriminating behaviours between genotypes. In the current study allogrooming and sniffing were rarely observed in Tph2-deficient mice and were only positively correlated to each other in control mice. These two behaviors are essential for social communication (Berg 2018 (silverman) doi: 10.1002/aur.1925) for the animal to build social knowledge (lee et al., 2019) and for maintenance of cohesion of group in mice (refs-allogrooming, scents and fear, see Weiss papers?). Sniffing is another important modality for communication which is used to establish the leadership position in animals (Wesson 2013). ~~When a subordinate is face-sniffed by a dominant animal the decrease of sniffs frequency by the subordinate has been proven to be an appeasing signal (ref). Sniffing is not only used to sample odors but also to communicate the animal’s status and have the capacity to rapidly influence conspecific behavior in a social-status dependent manner (Wesson 2013).~~ With a low sniffing count in Tph2-deficient mice (here and other Beiss 2015, Kane 2012) these animals did not use these important social cues for communication and behavioural adjustment. ~~In a recent study, the evaluation of the temporal microstructure of dyadic interactions during relationship formations in mice revealed “allogrooming” to be a critical indicator of established dominance (Lee 2019). While these behaviors would be similarly displayed by the unfamiliar individuals at first, once the relationship status would be resolved, both behaviors would be mainly only displayed by the dominant animal over the subordinate animal~~. In our study, such dynamical shifts in allogrooming and struggle networks were only seen in control groups. For allogrooming, the interaction strength between pairs of mice in Tph2-control networks statistically increased from day 2 to day 5 but not in Tph2-deficient groups. In other words, allogrooming was rarely displayed between pairs of Tph2-deficient mice over all days while it occurred more often and was mostly happening between same pairs in Tph2-control mice over time. In Tph2-deficient groups the lack of a dynamic shift in allogrooming is coherent with the delayed emergence of the alpha mouse and thus the diffuse nature of power distribution among the animals.

~~With this study we validated the use of this VBS housing-cage condition for analysis of the several layers, temporality and relationships of social and non-social behaviors in mice (from the individual to the group level).~~ Most of the social and everyday life impairments a patient suffers through are not easily measureable in laboratories or clinical settings and in pre-clinical research the use of ethological-like testing systems is still too often hindered by cumbersome or lack of adequate analytical tools’ availability. With this set-up and the corresponding analytical tools we developed for this project we could study the everyday-like symptomatology of our animal models. ~~Deadly aggression previously witnessed in this strain of mice when housed in their classical home cages was also prevented in the VBS despite aggression being still very prevalent in these young Tph2-deficient mice.~~ This highlights again the importance of considering the interaction between living arrangement and innate deficits one’s could carry (~~concept of pathology as a complex system~~). It is essential to consider that pathology only emerges as the limit of adaptation of an individual to internal/external constraints.

Mice lacking serotonin had poorer communication skills (i.e. sniffing and allogrooming), possibly poorer sensitivity to environmental cues (social and non-social), delayed short term memory formation and of social knowledge; a delay in hierarchical ranking and other social network dynamics. They lacked behavioural control of aggression but specifically in food-associated context (no de-escalation of aggression at feeder). In the Tph2-deficient mice, aggression appeared to be the visible part of an underlying complex interplay among various sensory, memory, communication, behavioural control and decision abilities. The role of the context of testing appeared essential and should be considered in further studies of the role of serotonin in the modulation of these intertwined functions in Tph2-deficient mice.

**Conclusion**

In this study, Tph2-deficient mice did not present all features of pathological aggression as it is classically defined. Beyond aggression, in their undisturbed housing conditions, they however revealed possessing a more subtle, complex and dynamic maladaptive phenotype. Tph2-deficient mice present a great potential to further investigate the role of serotonin in the coexisting expression of food related aggression control, short term memory formation, aspects of social competence and decision making which are all critical symptoms of most human mental disorders. Finally, with this project we have developed and made available to the scientific community a set of analytical tools allowing the study of complex freely expressed home cage phenomenon and to the dynamic and temporal characteristic of self-organizing group living mice.

**References (to do, zotero)**