

The influence of personality on territory use and fitness
in western house mice (*Mus musculus domesticus*)

Dissertation

Zur Erlangung des Doktorgrades

“Doctor rerum naturalium“

der Mathematisch-Naturwissenschaftlichen Fakultät

der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Rebecca Krebs-Wheaton

Max Planck Institute for Evolutionary Biology

Plön, October, 2017

First referee: Prof. Dr. Diethard Tautz

Second referee: Prof. Dr. Thomas Roeder

Date of oral examination: 20.12.2017

Table of Contents

Table of Contents	3
List of Figures	5
Acknowledgements	8
Zusammenfassung.....	9
Abstract	11
Contributions.....	13
1. Introduction	14
Personality and behavioural plasticity	15
Territory use	18
House mice and their social system	19
Scope of this thesis	19
2. Material and Methods	20
Study organism.....	20
Enclosure Experiment	21
Controls	23
Microsatellite genotyping	24
Population Structure and Parentage analysis	25
T haplotype.....	26
Population Statistics.....	27
Data Management	28
Behavioural Tests	28
“Quick tests” for all mice in each monitoring.....	28
3. Results	34

A. Population Development	34
B. Mate Choice	39
C. Territory Use	57
D. Behavioural Tests	78
<i>Population origin influenced behaviour</i>	83
E. Fitness in house mice	106
4. Discussion	115
Population development.....	115
Mate Choice	117
Territory use	117
Personality and behavioural plasticity in wild house mice.....	120
Differences in behaviour between two recently diverged populations	121
5. Conclusions and Outlook	125
6. References	127
7. Supplemental Information	134
Comments about technical issues in enclosures:.....	134
Digital supplement	140
Declaration/Erklärung	141
Curriculum vitae	142

List of Figures

1. Introduction

Figure 1. Examples of behavioural reaction norms. 17

2. Materials and Methods

Figure 2.1: Enclosure layout. 22

Figure 2.2: "Quick Tests". 28

Figure 2.3: "Longer Tests". 31

Table 2.1: Timeline. 24

Table 2.2: Polymorphic microsatellites. 25

3. Results

A. Population Development

Figure A.1: Structure Analysis of all individuals. 34

Figure A.2: Structure Analysis of founder individuals. 35

Figure A.3: Pure vs mixed individuals by enclosure. 39

Table A.1: Development of population in enclosures. 37

Table A.2: Number of individuals per generation in each enclosure. 38

B. Mate Choice

Figure B.1: Structure Analysis of all individuals. 40

Figure B.2: Proportion of mating events by enclosure. 41

Figure B.3: Number of successful of mating events for G1 individuals. 42

Figure B.4: Backcrosses between G1 and G0 mice. 43

Figure B.5: Reproductive success of G0 males by population. 46

Figure B.6: Reproductive success of G0 females by population. 48

Figure B.7: Reproductive success of G1 males by population background. 50

Figure B.8: Reproductive success of G1 females by population background. 52

Figure B.9: Reproductive success of G1 males by status. 54

Figure B.10: Reproductive success of G1 females by status. 56

C. Territorial Behaviour

Figure C.1: Locations used by mice. 58

Figure C.2: HRS maps. 59

Figure C.3: Relationship between territory scores. 60

Figure C.4: Mean AA scores for 7 monitoring events. 61

Figure C.5: Mean HRS scores for 7 monitoring events. 62

Figure C.6: Mean SI scores for 7 monitoring events 63

Figure C.7: G1 AA scores compared to parents. 65

Figure C.8: G1 HRS scores compared to parents. 66

Figure C.9: G1 SI scores compared to parents. 68

Figure C.10: AA scores of first generation offspring based on their status, maternal population and paternal population. 70

Figure C.11: The HRS scores of first generation offspring based status and maternal population. 71

Figure C.12: The SI scores of first generation offspring based on sex, status and maternal origin. 72

Figure C.13: Home range size of "Schuhmacher" and offspring. 75

Figure C.14: Home range size of "Verstappen" and offspring. 76

Figure C.15: Home range size of “Subaru” and offspring.	77
Figure C.16: Home range size of “Lasek” and offspring.	78
Table C.1: Generalized mixed effects model for AA scores of offspring to their parents.	69
Table C.2: Linear mixed effects model for HRS scores of first generation born in enclosures	71
Table C.3: Linear mixed effects model for SI scores of first generation born in enclosures.	72

D. Behavioural Tests

Figure D.1: Dark/Light Box test.	75
Figure D.2: Novel Object test.	76
Figure D.3: Open Field test.	77
Figure D.4. Elevated Plus Maze.	78
Figure D.5 Dark/Light Box test by population.	80
Figure D.6 Novel Object test by population	82
Figure D.7: Open Field test by population.	84
Figure D.8: Elevated plus maze by population	86
Figure D.9: Dark/Light box for control mice.	88
Figure D.10: Novel Object test for control mice.	89
Figure D.11: Open Field test for control mice	90
Figure D.12: Elevated Plus maze for control mice	90
Figure D.13: Correlation plot for repeatable measurements.	94
Figure D.14: Hierarchical Clustering of repeatable measurements.	95
Figure D.15: Behavioural scores.	96
Figure D.16: Creating behavioural scores by population.	99
Figure D.17: Behavioural reaction norms for Dark/Light Box and Startle test.	100
Figure D.18: Behavioural reaction norms for Novel Object test.	102
Figure D.19: Behavioural reaction norms for Open Field test.	103
Figure D.20: Behavioural reaction norms for Elevated Plus Maze.	104
Table D.1: Fixed effects model Dark/Light test (nose poke).	80
Table D.2: Fixed effects model Dark/Light test (mouse appears).	81
Table D.3: Fixed effects model Startle test (nose poke).	81
Table D.4: Fixed effects model Startle test (mouse appears).	81
Table D.5: Fixed effects for novel object behavioural measurements.	83
Table D.6: Fixed effects for the behavioural traits measured in open field test.	85
Table D.7: Fixed effects for behavioural measurements in the elevated plus maze.	86
Table D.8: Repeatability calculated using rptR.	91
Table D.9: Generalized linear models for jumps in quick tests.	98
Table D.10: Generalized linear models for bites in quick tests.	98
Table D.11: Generalized linear models for squeaks in quick tests.	98
Table D.12: Generalized linear models for tail jerks in quick tests.	99
Table D.13: Generalized linear model with random effects mouse looks out of shelter.	100
Table D.14: Generalized linear model with random effects mouse leaves shelter.	101
Table D.15: Generalized linear model with random effects mouse looks out of shelter post startle.	101
Table D.16: Generalized linear model with random effects mouse leaves shelter post startle.	101
Table D.17: Generalized linear model with random effects sniff latency.	102
Table D.18: Generalized linear model with random effects time spent with novel object.	102
Table D.19: Generalized linear model with random effects time spent at wall in open field.	103
Table D.20: Generalized linear model with random effects distance traveled in open field.	103
Table D.21: Generalized linear model with random effects time spent in dark arm elevated plus.	104

E. Fitness of house mice

Figure E.1: Offspring count by behavioural scores.	106
Figure E.2: Offspring count by territory scores.	109
Figure E.3: Offspring count by territory scores for males and females.	112
Table E.1: Fixed effects model for offspring count by behavioural scores.	107
Table E.2: Fixed effects model for offspring count by territorial scores.	108
Table E.3: Fixed effects model for offspring count by territorial scores of males and females.	111

Supplemental Information

Supplemental Figure S.1.A: Monitoring events in enclosure A.	135
Supplemental Figure S.1.B: Monitoring events in enclosure B.	136
Supplemental Figure S.2.A: Antenna reads enclosure A.	137
Supplemental Figure S.2.B: Antenna reads enclosure B.	138
Figure S3: Correlation matrix with Holm correction.	162
Description of Digital Supplement.	139

Acknowledgements

I would like to thank Diethard Tautz for making this possible, for supporting this project and guiding me through the process. Many thanks to Thomas Roeder and Thomas Schifflholz for joining my thesis committee and providing me with advice throughout the course of the project. Thank you to Diethard Tautz, Thomas Roeder, Oscar Puebla and Matthias Leippe for attending my PhD defense. Special thanks to Miriam Linnenbrink for her supervision, discussion and support throughout this project and the International Max Planck School and Max Planck Society for the funding and the opportunity they provided me with. Thank you Leslie Turner, for helping me with the final version of this thesis.

Many thanks to the entire mouse team, with special thanks to Christine Pfeifle, Heike Harre, Susanne Reinsch, Janine Wolf, Anika Jonas, Anastasia Vock and Maik Görtz-Sonnwald, who helped film the experiments and care for the mice. Thanks to Svenja Hoier, Nina Reinhardt, Jessica Pickard, Nina Hafer and Maryam Keshavarz for all the great discussions. Thank you to Harald Deiwick, for making the test boxes and mazes used during the experiment and customizing them for wild house mice. I would also like to thank Cornelia Burghardt and Heinke Buhtz for teaching me the lab work and Marie Vallier for teaching me to score microsatellites. Thank you to Frank Buschmann for the antennae and software that made this project possible and Patryk Stainski, for helping me keep them running. I would like to thank Barbara König and her group in the University of Zürich for all the advice they provided.

I would like to thank everyone who helped me get here, my first language professor Lothar Zimmermann for introducing me to CB/G literature, music and culture, and starting me on this path. I would like to thank Angela Osborn for helping me find the path again when I had lost it and ftw for keeping me on it until the end. I am extremely grateful to my friends and colleagues in Canada and Switzerland (especially Yvonne Zürcher) and my amazing office mates in Plön, Neva Skrabar, Stefan Dennenmoser and Sophie von Merten for scientific discussion, advice and support throughout the project. I would like to thank my family, especially my sister Hannah, for all their support throughout my studies. Special thanks to my grandmother Helena, for knowing the real value of education. I would also like to thank my partner Jan Wolff, for making my time in Kiel so wonderful and the Tigerlis for following me absolutely everywhere.

Zusammenfassung

Persönlichkeit bei Tieren ist definiert als ein konsistentes und stabiles Verhaltensmuster. Diese Definition beruht auf den Messungen tierischer Persönlichkeiten vieler unterschiedlicher Arten. Die Persönlichkeit von Tieren kann deren Fitness beeinflussen, da sie das tägliche Verhalten prägt. Das Verhalten eines Tieres hat direkten Einfluss auf den Kontakt mit Artgenossen und auf die Wahl potenzieller Partner. Die Verhaltenseigenschaften können die Zugangsmöglichkeiten zu Ressourcen beeinflussen, so z.B. die Wahl des Nistplatzes und die Erreichbarkeit von Futter. Des Weiteren können sie sogar die Wahrscheinlichkeit der potenziellen Begegnungen mit Raubtieren bestimmen. Laut zahlreicher Studien ist die Fitness von verschiedenen Persönlichkeitstypen stark durch deren Umwelt beeinflusst, z.B. die vorhandene Populationsdichte und die Verfügbarkeit von Ressourcen. Das heißt, dass die Fitness der verschiedenen Persönlichkeitstypen von Jahr zu Jahr differiert oder aber, dass die Persönlichkeitstypen, die nicht zu extrem sind, mehr Erfolg haben. Es wird allgemein angenommen, dass fluktuierende oder ausgleichende Selektion die Vielfalt der Persönlichkeitstypen in einer Population beibehält. Es kann also sein, dass die durchschnittlichen Verhaltenseigenschaften der verschiedenen Populationen anders sein können.

Wir möchten herausfinden, welche Rolle die Persönlichkeit auf die Fitness der westlichen Hausmaus spielt. Außerdem möchten wir die Verhaltensunterschiede zwischen zwei Populationen erforschen. Eine Population stammt aus Deutschland (Köln/Bonn) und die andere aus Frankreich (Zentralmassiv). Diese Populationen sind faszinierend, weil sie erst seit 3000 Jahren getrennt sind und sich dennoch bereits deren Genetik und Verhalten differenziert hat. Die Mäuse der beiden Populationen wurden für zehn Monate in einem gemeinsamen Gehege gehalten. Die Mäuse hatten die freie Wahl von Partnern und Revieren. Während dieses Zeitraums, wurden die Verhaltenseigenschaften der männlichen Mäuse der ersten Generation mehrmals im Labor durch Messungen bestimmt. Verschiedene Prüfungen sind dafür benutzt worden (novel object test, open field test, dark/light box test und elevated plus maze test). Die gemessenen Verhaltenseigenschaften wurden dann geclustert und die verwandten Eigenschaften

kombiniert. Die Hauptgruppen waren Aktivität, Ängstlichkeit und Neugier. Die Eigenschaften in jeder Gruppe wurden kombiniert, um ein "Behavioural Score" zu erschaffen. Während des Versuchs, war jede erwachsene Maus mit einem RFID-Chip markiert worden. Das hat die Messung der Aktivität und Reviergrößen ermöglicht. Die "Behavioural Scores", die Aktivität in den Gehegen und die Größe des Reviers waren mit der Anzahl des Nachwuchses korreliert worden. Schließlich möchten wir das von Montero *et al.* 2013 entdeckte Partnerwahlmuster prüfen, welches davon ausgeht, dass die Partnerwahl der Mäuse von den Vätern beeinflusst ist.

Wir haben entdeckt, dass der Durchschnitt der untersuchten Verhaltenseigenschaften bei den zwei Populationen (Deutsch und Französisch) unterschiedlich ist. Ausdrücklich unterschieden sie sich in der durchschnittlichen Aktivität und Neugier. Fitness und Persönlichkeit beziehen sich aufeinander. Im Fall der französischen Population haben aktivere Mäuse mehr Nachwuchs. Bei der deutschen Population gab es kein klaren Unterschied. Die Reviergröße war ebenfalls mit Fitness korreliert, wobei ein kleineres Revier mit einem Fitnessgewinn bei männlichen Mäusen verbunden war. Das Gegenteil war der Fall bei weiblichen Mäusen. Es war unmöglich das Partnerwahlmuster des Versuchs von Montero *et al.* 2013 zu replizieren, aber es könnte sein, dass unser Stichprobenumfang zu klein war, um das Partnerwahlmuster messen zu können.

Abstract

Consistent and stable patterns of behaviour are defined as animal personalities or behavioural syndromes. Animal personalities have been measured in a number of species, and it is now widely accepted that they exist in many taxa from insects to mammals. Personality traits may affect individual fitness, since they influence day to day behaviours, which directly affect the number of encounters they have with conspecifics, potential mates or predators. They may also influence access to food, nests and other resources. Several studies of the role personality plays in fitness have found that environment, including population density and resource availability, may influence the fitness of different personality types. This is why it is believed that these traits are under fluctuating or balancing selection, which are the two mechanisms proposed for the maintenance of diversity in personality types within populations. Furthermore, entire populations may differ overall in mean personality scores, due to differences in local environment.

We aim to investigate the role personality plays in the individual fitness of wild House mice (*Mus musculus domesticus*) in semi-natural enclosures. Furthermore, we wanted to look at overall difference in behaviour and fitness between two populations of house mice, one from Germany (Cologne/Bonn) and one from France (Massif Centrale). These populations are especially interesting, because even though they have only been separated for approximately 3000 years, they are already diverged genetically and behavioural, and exhibit a distinct mating pattern. Mice from both populations were kept for a 10 month period, in semi-natural enclosures where they were free to choose mates, establish hierarchies, build families and select housing sites. During this time, personality tests (Novel Object Test, Open Field Test, Dark/Light Box and Elevated Plus Maze) were carried out on founder males every four to five weeks and parentage analysis was performed to assess the influence of personality type on fitness (in terms of number of offspring). Behaviours were clustered to find relationships between measurements, and related measurements were classified into three groups, activity, anxiety-like behaviour and curiosity/risk-taking based on the measurements included, and combined to make one behavioural score for each category. In addition, the enclosures were equipped with RFID antennas and mice were tagged so that we were able to assess individual territory use and activity within the enclosures for all mice. Finally, we hoped to confirm the mating

pattern found by Montero *et al.* 2013, namely that mice born in the enclosures tended to choose mates with the same population origin as their fathers.

We found overall differences in behaviours between the French (MC/F) and German (CB/G) population in scores related to activity and curiosity (risk-taking behaviour). We also found that personality was related to fitness, and that there was an interaction between population origin and activity and anxiety-like behaviour, which seemed to affect the MC/F population more than the CB/G population. We also found that smaller home range size (HRS) was related to higher reproductive success in male mice but larger HRS was related to higher reproductive success in females. We did not replicate the mating pattern found by Montero *et al.* (2013), although this may be due to a smaller sample size, since values for paternal matching were approaching significance in some cases.

Contributions

This experiment was jointly planned by me and my supervisors Diethard Tautz and Miriam Linnenbrink. Mice in cages and enclosures were cared for by the entire mouse team, but especially Susanne Reinsch. Filming of enclosure mice was carried out under my supervision by the mouse team, Susanne Reinsch, Janine Wolf, Anastacia Vock, Annika Jonas, Christine Pfeifle, and Svenja Hoier. Monitoring of the enclosure mice were carried out by Heike Harre, Susanne Reinsch and myself. Videos scoring and lab work were carried out by me. Analysis of data was carried out by me, under guidance from my supervisors.

The influence of personality on territory use and individual fitness

1. Introduction

Behavioural mechanisms of pre-zygotic reproductive isolation may lead to population divergence and may over time result in speciation (Coyne and Orr, 2004). Mate choice can be an important mechanism of reproductive isolation, since assortative mate choice can lead to speciation, even when populations are sympatric or parapatric (Coyne and Orr, 2004; Dieckmann *et al.* 2004). Under these conditions, the development of different mating cues and preferences can lead to the separation of populations. In house mice, chemical (Potts *et al.* 1991; Potts and Penn, 1998; Malone *et al.* 2001, Beynon and Hurst, 2003) and auditory (Hammerschmidt *et al.* 2009; Musolf *et al.* 2010; Asaba *et al.* 2014; von Merten *et al.* 2014) cues are believed to play a role in mate choice and could influence mating patterns. One of the most famous examples of this occurs at a hybrid zone stretching from Denmark to the Caucasus between the two sub-species of the house mouse (*Mus musculus musculus* and *M.m. domesticus*). Asymmetric mate choice based on chemical cues has been proposed to serve as a reproductive barrier between subspecies (Smadja and Ganem, 2002) maintaining this hybrid zone. In addition, mating patterns have also been measured between house mice from the same subspecies (*M.m. domesticus*) (Montero *et al.* 2013). In this case, the house mouse mating system was studied using mice from two allopatric populations of *M.m. domesticus*, one from Germany (Cologne-Bonn area: CB/G) and the other from France (Massif Central area: MC/F). Mice from these populations were placed in semi-natural enclosures together and mate choice was studied via genotyping the offspring over the course of five to seven months (Montero *et al.* 2013). These populations, which have only been separated for approximately 3000 years (Cucchi *et al.* 2005) already show considerable genetic differentiation (Ihle *et al.* 2006; Teschke *et al.* 2008; Staubach *et al.* 2012), as well as in the expression of genes associated with behavioural traits (Bryk and Tautz, 2013; Lorenc *et al.* 2014). Montero *et al.* (2013) found that founder mice in the enclosures mated assortatively, but only if they were familiar with individuals from their own population first. They also found that mice from the second generation showed a specific mating pattern, influenced heavily by the paternal population origin.

Personality and behavioural plasticity

Mate choice may also be influenced by animal personality. Personality in non-human animals is defined as patterns of correlated behaviours expressed by individuals, which are repeatable across time and contexts (Gosling, 2001; Sih *et al.* 2004). Gosling (2001) provides a summary table of personality measurements from a variety of species published in the last century (since then there have been many more), focused mainly on primates and other mammals, but including examples of behavioural studies in fish, birds, reptiles, amphibians and invertebrates, where behavioural traits were measured and found to be consistent across contexts. These studies usually focus on activity, the bold/shy axis, exploration, aggression and sociability (Reale *et al.* 2007). Personality traits are heritable (Van Oortmerssen and Bakker 1981; Benus *et al.* 1991; Koolhaas *et al.* 1999; Dingemanse *et al.* 2002; Veenema *et al.* 2003b, van Oers and Mueller, 2011), although there is still a lack of information on exact mechanisms of inheritance. In humans, personality traits are believed to be complex traits and are probably influenced by a number of genes (Knopik *et al.* 2016). In human twin studies, heritabilities of the OCEAN traits (openness, conscientiousness, extraversion, agreeableness and neuroticism) were 61%, 44%, 53%, 41% and 41% respectively and genetic influence was mostly non-additive (Jang *et al.* 1996). In animal models, risk-taking behaviour is estimated to be 19% heritable, while exploratory behaviour was around 54% heritable in *Parus major* (van Oers *et al.* 2004). In laboratory mice, anxiety-like behaviour is estimated to be between 30-59% heritable, although quantitative trait loci (QTL) analysis could only explain 3.5-26.5% of the variance in behaviour (Gerschenfeld and Paul, 1997). Behavioural patterns may also be influenced by the environment, which means they can be subject to plasticity due to experience (Sih *et al.* 2004).

Consistent differences in animal behaviour could also contribute to reproductive isolation and speciation (Ingley and Johnson, 2014). Theoretically, differences in personality could limit the amount of contacts an individual has with conspecifics and potential mates. This could be due to temporal or foraging behaviour or differences in habitat use (Ingley and Johnson, 2014). This could not only influence contact with potential partners and intraspecific competition for resources, but it may also increase or decrease the likelihood on confronting predators. In laboratory studies between two species of fire-bellied toads (*Bombina bombina*), different behavioural types (the more active *B. variegata*) are more susceptible to predation than the less active *B. bombina*, which may help to maintain the hybrid zone (Kruuk and Gilchrist, 1997).

Changes in average individual behaviour resulting from environmental conditions or experience are defined as “behavioural plasticity”. An individual’s behavioural plasticity demonstrates the range in which an individual can modify its behaviour depending on context and could be a kind of personality trait (Dingemanse *et al.* 2010). Dingemanse *et al.* (2010) published a summary table of behavioural traits, contexts and variation in environments by species. Two examples of small mammals, the Golden Hamster (*Mesocricetus auratus*) and the Eastern Chipmunks (*Tamias striatus*) showed behavioural plasticity resulting from the number of trials (reduction of novelty) (David *et al.* 2004, Martin and Reale, 2008). In the first example, hamsters which were less aggressive adapted faster to a bar pressing task, compared to more aggressive individuals (David *et al.* 2004), while Eastern Chipmunks habituated to experiments over time (Martin and Reale, 2008).

The degree to which an individual can adapt its behaviour according to context could present a considerable fitness advantage to individuals which can modify behaviour under different conditions, whereby individuals which rigidly conform to one behavioural type could be very successful in one environment but not in others. Plasticity in behavioural traits is often represented by calculating behavioural reaction norms which show the variation in behaviour by context for individuals (Dingemanse *et al.* 2010). Figure 1 (adapted from Schuett *et al.* 2009) shows examples of differences in plasticity according to two different contexts. Context can be anything from changes in social environment (e.g. population density), seasons, physiological state or time (Dingemanse *et al.* 2010). Habituation to experimental set ups could also be demonstrated using behavioural reaction norms (David *et al.* 2004, Martin and Reale, 2008).

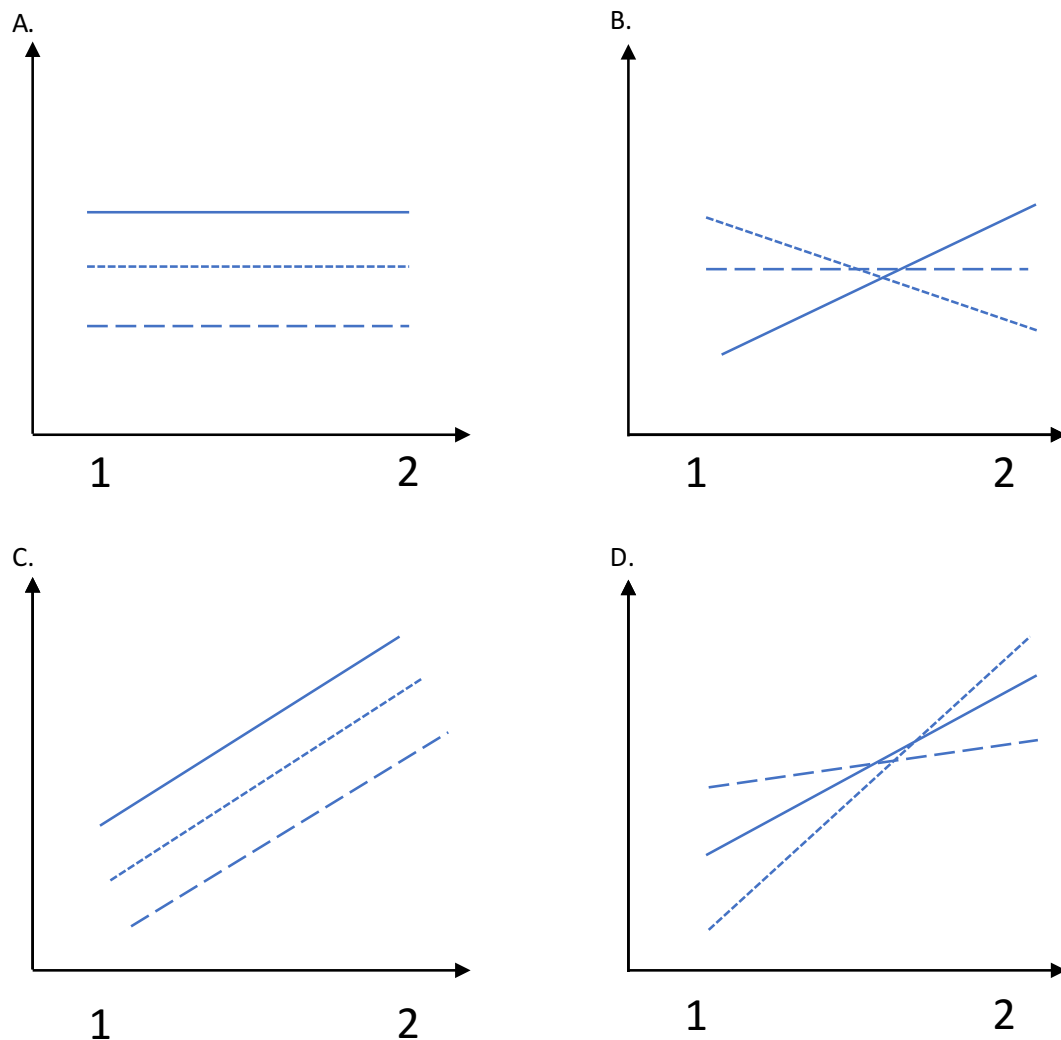


Figure 1: Examples of behavioural reaction norms. Adapted from Schuett *et al.* (2009) depicting three individuals in two different contexts. A. no plasticity in a behavioural trait B. no overall inter-individual differences with no consistency, C. inter-individual differences in a behavioural trait and variation but the same rank order and D. inter-individual differences with slope with different slope (no behavioural consistency) between contexts.

As stated above, consistent behavioural patterns (personality) may influence individual fitness because under certain conditions they may affect individual survival and reproductive success (Ingley and Johnson, 2014). It is believed that these traits are under “fluctuating selection” (Dingemanse and Reale, 2013) allowing them to remain variable within populations. There are two main theories to explain how this may work; the first, changing environments, population density and food availability may favour one behavioural type over another under certain conditions (Reale and Festa-Bianchet, 2003; Dingemanse *et al.* 2004; Both *et al.* 2005; Le Coeur *et al.* 2015; Nicolaus *et al.* 2016). The second suggests that benefits

from adopting one strategy or another may be balanced out by costs incurred by the same strategy, thus making all behavioural types equally fit (Sih, *et al.* 2004). In both cases, the explanations only apply if traits are not highly plastic and if individuals cannot simply adjust their behaviour to accommodate changing circumstances.

Recent studies into the role personality plays in fitness have demonstrated that different personality types have different benefits under different circumstances. For example, Great Tits (*Parus major*) which explore their environment faster produced more offspring in years where the population density was lower (Nicolaus *et al.* 2016) and long term studies of Siberian Chipmunks (*Tamias sibiricus*) showed that bolder individuals had higher reproductive success in years when food was scarce (Le Coeur *et al.* 2015). This study took six years, and they found the results to be repeatable.

Territory use

Territoriality is defined as the defense of a space or home range (Nice, 1941), usually by dominant individuals defending a territory or resource against subordinates (Kaufmann, 1983). The study of territoriality in wild house mice began with Peter Crowcroft's experiments in semi-natural enclosures in 1966 ("Mice all over", Crowcroft 1966). Since then, many studies have shown that mice are territorial animals, using urine marks for defining their territory and attract mates; these marks provide information about their social status (dominant or subordinate) (Hurst, 1990) and about individual identity through major urinary proteins (MUPs) (Hurst *et al.* 1998). Personality may influence territorial behaviour. In Great Tits for example, exploration speed determined the individual response to perceived intrusions on territories, whereby individuals with lower exploration scores confronted "intruders" (actually a loud speaker playing bird song) less often (Amy *et al.* 2010). But, there are no studies focused on this topic in house mice. Since urine marking is used to mark territories and carries information about the social and sexual status of the "marker" (Oakeshott, 1974; Bronson, 1979; Potts *et al.* 1991 and Montero *et al.* 2013), it is very possible that territorial behaviour also influences the fitness of a mouse, by limiting access to resources.

House mice and their social system

Mice reproduce rapidly (three generations per year) and produce on average five to eight pups per litter (Berry and Jakobson, 1971). Wild juvenile mice disperse and form new colonies (Bronson, 1979). House mouse populations are believed to be structured (Singleton and Hay, 1983); made up of dominant and subdominant individuals and there is low gene flow between these subpopulations (Bronson, 1979; Singleton and Hay 1983). Studies have shown that mice are territorial animals, using urine marks to defend their territory, attract mates and broadcast their social status (Hurst, 1990). Mice also nest communally, often in family groups (Montero *et al.* 2013), but female mice are also known to choose female nesting partners (Manning *et al.* 1995) that are not necessarily relatives (Weidt, *et al.* 2014). Finally, it is known that female mice are polyandrous (and male mice are polygynous) and 30% of litters are of multiple paternity (Montero *et al.* 2013).

Scope of this thesis

Studying personality and territoriality in wild house mice (*M.m. domesticus*) is particularly promising due to the vast knowledge of mouse genetics and the role laboratory mice play in medical and behavioural research. The great advantage of studying wild house mice is the natural behavior and genetic variation, which is not present in laboratory mice. The relatively fast reproductive cycles of mice in general allows for the study of fitness in several generations in a shorter time.

This study addresses three different aspects of natural behavior: 1) mate choice behavior, 2) competitive potential/territoriality and 3) personality traits in wild mice and I aim to elucidate the relationship between personality and territoriality and its influence on individuals fitness in two recently diverged populations of the Western house mouse (*M.m. domesticus*). We therefore have designed a semi-natural enclosure experiment based on the previous study by Montero *et al.* (2013), using a Aniloc Antenna System (RFID ring antennae, FBI Science GmbH) used in other semi-natural enclosure experiments (Koenig *et al.* 2015). We used three semi-natural enclosures (rooms), which were populated by wild-derived house mice from the MC/F and CB/G populations tagged with RFID transponders. These populations were chosen because they have shown behavioural differences in previous experiments (Montero *et al.* 2013; von Merten *et al.* 2014; Bachelor Thesis, Nina Reinhardt).

In this enclosure environment, mice were mainly undisturbed and could establish hierarchies, form social bonds and mate with a self-chosen partner. By this setup it was possible to investigate mate choice behavior within and between the two study populations and perform parentage analyses at the end of the experiment, which was continued until the second-generation mice in the experiment could choose mates to compare results of potential mating patterns to that of Montero *et al.* (2013). Further, founder male mice were tested at regular intervals in controlled behavioural experiments, adapted from medical research, throughout the course of the experiment to measure potential personality traits and classify them into types. These results were combined with the territory use data provided by the antennae. Thus, we were able to make inferences about the influence of personality and territory use on the fitness of individuals in terms of number of offspring and compare the success of different behavioral strategies.

2. Material and Methods

Study organism

Experimental mice were progeny of mice that originated from wild populations sampled in the Massif Central region of France (MC) (44°15'N–44°30'N, 2°45'E–3°15'E) and the Cologne/Bonn region of Germany (CB) (50°45'N–51°N, 6°45'E–7°E) in 2004 and 2005. In the animal facilities at the Max-Planck-Institute for Evolutionary Biology in Plön, mice were kept under an outbreeding regime to maintain genetic diversity within the populations. In general, all mice were kept in standard lab cages (Type III, Bioscape, Germany), and were weaned at the age of 28 days. Males were housed together with brothers, or in individual cages. Females were housed in sister groups, to a maximum of 5 mice per cage. Woodchip bedding (Rettenmaier, Germany) was used to line cages. Enrichment, including: wood wool, toilet paper, egg cartons and a spinning wheel (Plexx, Netherland), was provided in each cage. Mice were fed Standard Diet 1324 (Altromin, Germany) and provided water *ad libitum*. Housing prior to experiments was approximately 20–24°C, 50–65% humidity and maintained on a 12:12 light-dark schedule with lights on at 7 am. Mice used for the controls were housed individually in standard lab cages (Type III, Bioscape, Germany), and kept under the same conditions and light-dark cycle as they were prior to the experiment.

Enclosure Experiment

Three replicates were run in parallel in two rooms measuring 19.6 square meters and one measuring 17.6 square meters. A total of 108 mice started in the semi-natural enclosures, each enclosure started with 36 mice, half from the MC/F and CB/G population at an equal sex ratio (later on called “founder mice” and G0). All mice were between 3 to 6 months old at the beginning of the experiment and were tagged with RFID transponders (ISO Transponders, Planet ID) when they weighed 12 grams and ear clips were taken for later parentage assignment (stored in 70% Ethanol at +4°C). Some founder mice were related in each enclosure, but no more than one sibling was placed in each enclosure and only a sibling which was not from the same litter. The enclosure (Figure 2.1) was lined with Tapvei (Estonia) aspen bedding and equipped with standard cinder blocks, small wooden climbing walls (40 cm high) and rope made with natural fibers for enrichment. Each enclosure was equipped with twelve nest boxes; each stuffed with wood wool and toilet paper. Each house had one transparent entry pipe, 5.5 cm diameter. Other shelters included six 50 cm long opaque grey pipes. All shelters had 2 Aniloc (FBI Science GmbH, Viersen) antennae attached at each end to track the direction of the mouse entering and exiting the shelter. Additionally, 12 antennae were attached to water bottles at the front of the enclosure for tracking access to resources. 12 food dishes were also positioned in the front of the enclosure and food and water were provided *ad libitum* (Figure 2.1). The software “Olcus” (FBI Science GmbH, Viersen) was used to track the mice as they moved through the antennae. Temperature and humidity in the enclosures varied over the year based on natural fluctuations in the ambient environment. The light-dark cycle was 12:12 hours.

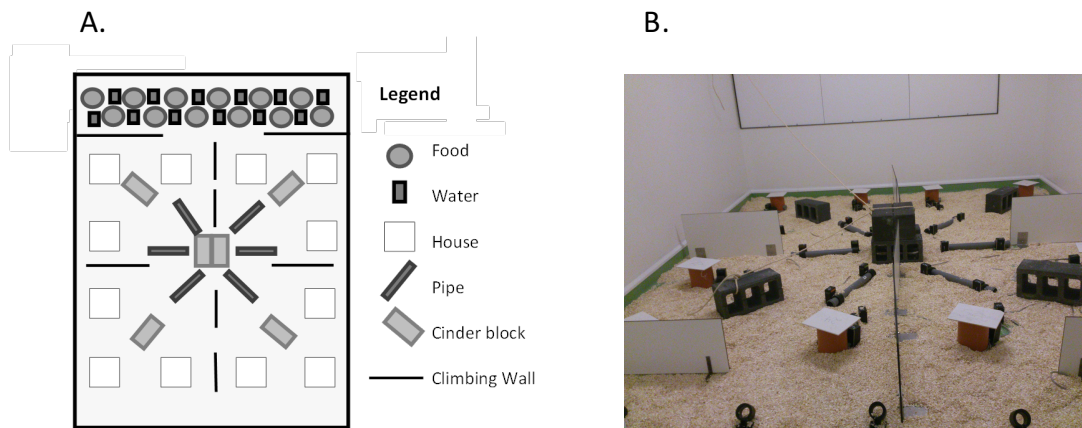


Figure 2.1: Enclosure layout. A. Drawing of enclosure. B. Photograph of an enclosure C (room 14). Houses were spaced between 90 cm to 110 cm apart. Water bottles and food sites were placed at 50 cm intervals in the front of the enclosure. Climbing ropes were provided for enrichment.

The enclosures were checked daily for food and water. More detailed checks of the enclosures were performed at 4-5 week intervals (8 times welfare checks, 7 times with behavioural tests). During these checks, all mice were taken out of the semi-natural enclosures, in traps or by closing the houses. All mice were weighed and subjected to the quick behaviour tests (see below). Unmarked mice were given transponders and an ear clip was taken for later genotyping. Mice caught in traps were kept in an arena with food Standard Diet 1324 (Altromin, Germany), water and shelter. Mice caught in houses that were not used in further behavioural tests were temporarily housed in standard lab cages (Type III, Bioscience, Germany) and provided food, water and shelter, before being placed in a fresh house, and returned to the location in which their previous house was found. Mothers with young pups were left in their house to ease stress and thus were not tested in behavioural tests. Mice used in the long behaviour tests included all founder male mice and 10 of the next generation mice. They were individually moved to standard lab cages (Type II, Bioscience, Germany), with food, water and shelters to await the longer behavioural tests (see Material and Methods, and Figure 2.3). While monitoring enclosures, welfare checks were carried out and injured mice were removed from the experiment and sacrificed using CO₂/O₂ gas mixture. At the end of each monitoring, mice were returned to the enclosure in their houses or released at the entrance if they were caught in traps.

Controls

Controls were made to control for habituation to the tests due to handling. 24 mice were tested in all behavioural tests (see below); this included 12 mice from the CB/G population and 12 mice from the CB/G population. Furthermore, half of the mice from each population were male and half were female. All control mice were kept in the same enclosure and mixed together on the same rack regardless of experimental condition. These mice were kept individually in standard lab cages and were not released in the semi-natural enclosures. The control group is divided into groups of 12, one group which was tested regularly following the same monitoring schedule (7 times, at 5 week intervals) as the mice in the enclosures; the other was tested in all behavioural tests only twice, once at the beginning in April, and once in October, 7 months after, reflecting the time frame of the semi-natural enclosure experiment.

Timeline

Enclosure experiments were run between April 2015 and March 2016. Table 2.1 shows the timeline for experiment and controls.

Table 2.1: Schedule for the enclosure experiment. Behavioural tests mentioned are described more detailed below, quick tests and long tests.

April 2nd 2015	Control 1 and Control 2
April 14th-20th 2015	Mice placed in Enclosures
April 27th-May 8th 2015	First Monitoring (Dark/light test and Elevated Plus Maze, quick tests) Control 1
May 19th -May 27th 2015	Second Monitoring (Open Field test and Novel Object Test, quick tests)
June 11th 2015	Control 1
June 23rd-June 30th 2015	Third Monitoring (Dark/light test and Elevated Plus Maze, quick tests)
July 16th 2015	Control 1
July 29th- August 4th 2015	Fourth Monitoring (Open Field test and Novel Object Test, quick tests)
August 27th 2015	Control 1
September 1st-September 8th 2015	Fifth Monitoring (Dark/light test and Elevated Plus Maze, quick tests) Control 1
October 05th-October 13th, 2015	Sixth Monitoring (Open Field test and Novel Object Test, quick tests) Control 1 and Control 2
November 17th-November 24th, 2015	Seventh Monitoring (All Tests, some F1 mice)
February 9th -11th, 2016	Eighth Monitoring: Welfare Check
March 9th-11th, 2016	End of Experiment and Clean-up

Microsatellite genotyping

At the end of all experiments, mice were sacrificed using a CO₂/O₂ mixture and cervical dislocation. Tissue samples were taken from all mice that had not been previously sampled. They were included in further parentage analysis. Ear clips were taken from all mice in the enclosure experiments and stored in 70% alcohol at +4°C. DNA was extracted using either salt extraction or DNeasy 96 Blood and Tissue Kit (QIAGEN) following the protocol for animal tissue, including extended centrifugation and decreased amounts of AE buffer at the final step (30 µl). DNA samples were typed for up to 13 microsatellite loci

(Table 2.2) using the standard protocol for the QIAGEN Multiplex PCR kit. Alleles were called using Geneious R8.1.8 (Biomatters Ltd).

Table 2.2: Polymorphic microsatellites used in Montero *et al.* 2013 in parentage analysis of potential founder mice (described in Teschke *et al.* 2008).

Locus	Number of alleles	Number of individuals typed	Observed heterozygosity	Expected heterozygosity	Polymorphic information index	Estimated nu allele frequencies
Chr16_21	11	103	0.79	0.86	0.84	0.04
Chr3_24	14	103	0.83	0.89	0.87	0.03
Chr12_05	11	117	0.82	0.86	0.85	0.02
Chr01_25	11	115	0.84	0.87	0.85	0.02
Chr10_45	14	116	0.66	0.84	0.82	0.12
Chr05_45	12	118	0.65	0.75	0.72	0.07
Chr13_22	10	118	0.67	0.80	0.77	0.09
Chr17_09	16	118	0.75	0.84	0.81	0.05
Chr19_08	12	119	0.77	0.85	0.83	0.06
Chr09_20	13	120	0.78	0.87	0.85	0.05
Chr14_16	14	119	0.66	0.87	0.85	0.14
Chr01_23	12	119	0.73	0.82	0.80	0.06
Chr02_02	11	117	0.73	0.83	0.80	0.07

Population Structure and Parentage analysis

Structure 2.3.4 (Pritchard, Stephens and Donnelly, 2000) was used to verify population structure using Bayesian clustering. Since mice from two populations went into the experiment, K=2 was expected, two runs were made with K=2 and two runs with K=3 (1,000,000 burn-in length and 1,000,000 MCMC simulations) (as in Hardouin *et al.* 2015) to verify that only two populations are found. This allowed us to verify the hybrid and pure offspring of the MC/F and CB/G populations. MS Analyser (MSA) (Dieringer *et*

al. 2003) was used to calculate pairwise distances between individuals and POSA table (proportion of shared alleles). MEGA6 (Kumar *et al.* 2016) was used to construct Neighbour Joining trees and visualize alleles shared between the two populations following Ihle *et al.* (2006).

Parentage analysis was carried out to measure the reproductive fitness of individual mice. Cervus 3.0.7 (Kalinowski, 2007) uses multiple autosomal loci to assign potential parents using pairwise likelihood to assign pairs of parents to offspring. Offspring were assigned to parents based on time periods. Only mice which had been alive in the last monitoring and mice alive during the current monitoring were considered as potential parents for the current monitoring. Potential offspring were considered if they were born during the current monitoring period and estimated age (14-21 days old). Log-likelihood is calculated for each candidate true parent, and then compared to the log-likelihood (LOD) of a random potential parent. The difference between these two values is referred to as delta. To determine the critical value of delta for a given data set, simulations were run for 10,000 assuming 95% of potential parents were sampled. Out of 13 loci, a minimum number of typed loci of 7 and a maximum number of mismatches of 2 were used. Strict confidence intervals were set at 95% and relaxed confidence intervals were set at 80%. Cervus was used to calculate the number of loci (k), observed heterozygosity (H_{obs}) and expected heterozygosity (H_{exp}), polymorphic information content (PIC), a value between 0 and 1 which represents the diversity of a DNA segment within a population, and the estimate of null alleles (F_{null}) (see Table 2.2). All assigned parents were verified manually based on antenna information and the location where potential mothers and pups were found during monitoring events and the number of potential fathers assigned per litter (e.g. too many fathers per litter would be dubious, Montero *et al.* 2013). Allele frequencies were calculated for each 8 points in time, corresponding to the 8 monitoring periods, based on the population that was currently in each enclosure (enclosures were calculated separately).

T haplotype

The t haplotype is a selfish genetic element on chromosome 17 that affects the fitness of the carriers. Males which are t/wt produce sperm with a higher likelihood of fertilizing eggs than wt/wt sperm (Lyon, 2003); however, the homozygous t/t genotype is lethal. Females carrying the t haplotype avoid fertilization by t/wt males (Manser *et al.* 2015). To ensure that the presence of the t haplotype did not skew the results for the mate choice test, the parents of the mice used in the experiment were tested for its presence.

Mice were tested for Hba-4ps and Tcp-1 using the same method as in Montero *et al.* 2013 (Schimenti & Hammer 1990; Morita *et al.* 1993). Amplifications of the following sequences Hba4ps_F: 5'-gagtgcactgcatgcccacaagctgtg-3' and Hba4ps_R: gagctgtggagacaggaagggtcagtg-3' (Schimenti & Hammer 1990) and Tcp1_F: 5' - gacaatcatagccttctcag-3' and Tcp1_R: 5'-gcagtgttatctttcactgg-3' (Planchart *et al.* 2000) marked with FAM labels on the 5-prime end. Fragments were amplified using 2-5 ng of DNA in 5 µl reactions using a multiplex PCR kit (Quiagen). Two separate PCR protocols were used in ABI Fastcyclers. For Hba-4ps: 95°C for 15 minutes, 28 cycles (30 seconds) at 95°C, 1:30 minutes of annealing time at 66°C, 72°C for 1:30 and the final extension time 10:00 minutes at 72°C. For Tcp-1: 95°C for 15 minutes, 28 cycles at 95°C (30 seconds), 1:30 of annealing time at 58°C, 1:30 minutes at 72°C and final extension time of 10:00 minutes extension time at 72°C. After amplification, samples were diluted with 20 µl of HPCL water and incubated at enclosure temperature for 30 minutes. 1 µl of the PCR products was added to 10 µl HiDi formamide + 0.01 µl ROX 500 size standard (ABI) per well. Plates were then denatured in thermocycler, 2 minutes at 90°C and 5 minutes at 20°C. Fragments were run in ABI sequencer and analysed using Geneious R8.1.8 (Biomatters Ltd). Mice which were t/wt show fragment sizes for Hba-4ps at 214bp/198bps and TCP1 600bp/425bps. The presence of both of the longer indicated the presence of the t haplotype. Some parents of the founder mice had both Hba-4ps at 214bp/198bps but no mice carried the longer fragment at TCP1, suggesting no mice carried the t haplotype.

Population Statistics

Statistical analysis was done using R 3.0.3 (R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>) and Microsoft Excel (2007). For comparisons of genotypes within each enclosure, chi-square test and two-tailed Kruskal test were used in R 3.0.3. Significance level was set as the standard $p \leq 0.05$. Graphical presentations were made in R, using standard software and "ggplot2" (Wickham, 2009).

Data Management

Data was managed using databases made in Microsoft Access 2007 and Microsoft Excel for Mac worksheets. The database includes all individual mice, transponder numbers, sex, condition and population (for the original mice). Tables were made for each enclosure and these were divided into monitoring events (M1-M8). One final table was made for the last monitoring, when all mice were sacrificed and the enclosures were taken apart (Final). Raw values from behavioural tests are also stored in excel sheets by test and PDFs.

Behavioural Tests

“Quick tests” for all mice in each monitoring

In each monitoring, every mouse (males and females) was tested in two “Quick tests” after being weighed. The first test consisted of one minute sitting in a clear PVC transfer pipe, in which the number of jumps were counted. The second comprises one minute being held at the base of the tail on a flat surface, in which bites, squeaks and escape attempts are counted. For images of these tests, see Figure 2.2.

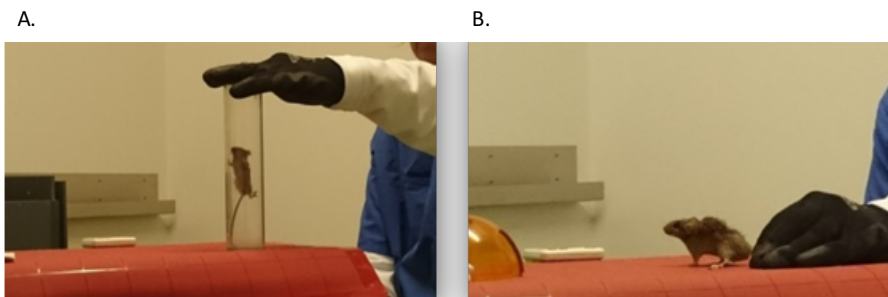


Figure 2.2: “Quick Tests”. Every mouse was tested in both quick tests for activity and aggression. First, the mice were placed in a clear plastic pipe (A) and the number of jumps was counted. Second, (B) the mice were held by the tail and the number of bites and squeaks were counted.

Long behaviour tests

During the experiment, there were 7 monitoring events in which mice were filmed, occurring at 4-5 week intervals. Founder males and 10 individuals from the first generation, born in the enclosures were

tested in the longer behavioural tests. Females were excluded for these tests to reduce stress and therefore possible stress related effects on pregnancy and offspring. During each monitoring event, the longer behavioural tests were alternated between two sets so that the mice were not excluded from their house and family for too long: set one (dark/light test and elevated plus maze) and two (open field and novel object tests), taking between 15 and 20 minutes per mouse tested. These tests were adapted from tests commonly used in biomedical research on laboratory mice and tests carried out on laboratory mice, wild common voles and striped mice (Archer, 1973; Gosling 2001; Herde and Eccard, 2013; Yuen *et al.* 2015). These experiments were filmed using TSE cameras and tracked and scored using VideoMot2 (TSE Industries, Inc). See Figure 2.3.

Elevated Plus Maze

The focal mouse could explore a modified elevated plus maze, with clear Plexiglas arms and a lid (Figure 3A). This modification with the lid was necessary as in the trial phase of these tests wild mice showed an ability to easily escape the otherwise commonly used setup. Each arm was 50 cm long with a 10 centimeters neutral area in the center where the arms crossed. Two of the arms were made of clear Plexiglas and two were made of grey PVC. The floor was made of white PVC. Each mouse could explore the maze for five minutes (adapted from Holmes *et al.* 2000). During this experiment, the time spent in the dark and light arms were measured, as well as the speed and distance travelled. At the end of the experiment the setup was cleaned with 30% Ethanol.

Open Field

The focal mouse could explore a 60x60 cm test apparatus (Figure 3B) for 5 minutes (Wilson *et al.* 1976; Reale *et al.* 2007; Yuen *et al.* 2015). The speed of the mouse, the distance travelled and time spent within 10 centimeters off the wall vs in the central area were measured. This test was directly followed up by the Novel Object Test.

Novel Object Test

In the same test apparatus as the open field test, a Lego toy was placed in the center (Figure 3C). The time of the novel object was explored, and then the time spent sniffing the new object was measured

over a period of 5 minutes (Birke and Archer 1983; Verbeek *et al.*1994, Yuen *et al.* 2015). At the end of the experiment the setup was cleaned with 30% Ethanol.

Dark/Light Box

The focal mouse was placed in a test apparatus containing a small dark shelter with two exits (Figure 3D). During the first five minutes, the time until the mouse pokes its nose out of the shelter and the first time the tail is visible was recorded. At five minutes, a set of keys was dropped next to the test apparatus, and the second part of the experiment began. The time it took for the mouse to first look out and when the entire mouse was visible was measured. If mice did not come out at all, the time was set to be 600 seconds. This test was adapted from tests used in lab mice and common voles (Young and Johnson, 1991; Herde and Eccard, 2013). At the end of the experiment the setup was cleaned with 30% Ethanol.

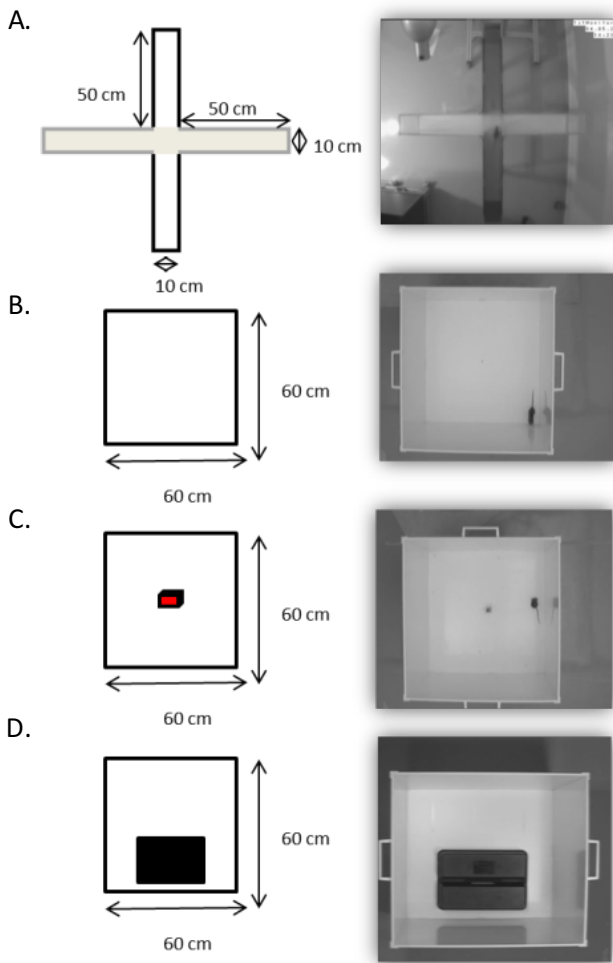


Figure 2.3: Setups for the “Longer Tests”. The figure shows diagrams and pictures of experimental setups including measurements of the setup. (A) the elevated plus maze, where the grey arms represent the open arms and the closed arms are depicted in black. The depth of the side walls is 40 cm. (B) the open field test, which serves as the base for the novel object test (C) and the dark/light test (D).

Calculating Home Range Size and Activity

Mice were tagged with transponders (ISO FD-X Transponders, Planet ID) and mice were tracked using AniLoc Antennae (FBI Science, GmbH, Viersen). Location change data gathered using Olcus software (FBI Science, GmbH, Viersen) was used to estimate activity in the arena (AA) based on antenna reads and home range size (HRS) based on locations where the mice spent 95% of their time. We used R version 3.3.2 (R Core Team, 2016) base program and the packages “sp” (Roger *et al.* 2013) to create spatial objects and “adehabitatHR” (Calenge, 2006) to calculate minimum convex polygons to describe HRS.

Site Selectivity

A Selectivity Index (SI) was calculated for each individual by calculating the standard deviation score (SD) for each mouse for all possible sites between each monitoring period and dividing this score by the maximum possible standard deviation given that there are 30 possibilities ($SD_{max} = 0.1826$) (Linnenbrink and von Merten, 2017).

Heritability or Imprinting of Territorial Behaviour

To estimate heritability of AA, HRS and SI the monthly scores of G0 parents were averaged to create one score for AA, HRS and SI for the entire 8-month period. The AA, SI and HRS scores for the offspring were calculated in the same way. AA and HRS data were transformed using a log₁₀ transformation. SI scores were square root transformed to normalize residuals in the selectivity model. The AA score for offspring was compared to the AA and HRS scores for the mother, the father and the average of the two parents and a Pearson correlation test was used to test the relationship between these scores. Furthermore, the R package “nlme” (Pinhiero *et al.* 2017) was used to create random effects models to exclude external effects influencing correlations.

Statistical Analysis

Statistics were carried out using R 3.3.3 and R 3.3.2 (R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>). In the novel object test, speed and distance were normally distributed; all other measurements were not and for these non-parametric statistics were used. Generalized linear models to measure habituation and the effect population had on behaviours were carried out using “lme4” package (Bates *et al.* 2015). “Glmer” was used for data that were not normally distributed (poisson family for count data). For data that were not normally distributed and not count data, data were transformed by square root for left skewed data and log₁₀ for data skewed to the right. Furthermore, random effects models were made using the R package “nlme” (Pinhiero *et al.* 2016). Models were made using the minimal suitable model determined by gradually removing variables from the full model and selecting the model with the lowest AIC scores. Figures were made using “ggplot2” (Wickham, 2009).

To determine whether individual behavioural measurements are correlated, a Spearman correlation matrix was made. P-values were corrected using the Holm method. Behaviours were clustered using the protocol from (Herde and Eccard, 2013). An hierarchical cluster function was used from the R package "cluster" (Maechler *et al.* 2016), specifically "agnes", to determine the relationship between the measurements. All measurements were clustered using Manhattan clustering with complete linkage (Gyuris *et al.* 2011; Tremmel and Müller, 2013; Herde and Eccard, 2013). The resulting dendrogram is shown in the Results section (Clustering). The dendrogram shows all measured variables which were moderately repeatable and at what level two variables are joined ('height'). The height where the variables are joined shows the distance between two clusters (e.g. the shorter the distance the more similar the variables).

Repeatability was calculated using "rptR" package (Naganawa *et al.* 2010). For data which were normally distributed (Novel Object: speed and distance), the function "rpt.aov" was used or for a Gaussian distribution. For non-parametric measurements, "rpt.glmm" was used, with a poisson distribution and SI scores used a proportional model.

Principle component analysis was made using the "psych" package (Revelle, 2016), in R functions "factor.pa" and "principle" to assign behavioural scores for anxiety, curiosity and activity (based on the variables which clustered in the dendrogram). Behavioural scores were later compared to the individual fitness of mice based on the results of the parentage analysis. Spearman rank correlation was used to determine the relationship between fitness and behavioural score. Holm correction was used to adjust p-values for multiple measurements.

3. Results

A. Population Development

Each enclosure started off in mid-April with 36 individuals, half male and female and half from each population. To differentiate between population of origin 13 microsatellite loci were used (as in Montero *et al.* 2013). A neighbour joining tree was constructed from the initial population (Mega 6, Figure A.1) to verify population origin. Three CB/G individuals were assigned to the wrong population; this could be due to missing microsatellite data for these individuals (typed at 6, in one case 5). 6 was the minimum threshold used during the parentage analysis and this mouse was excluded.

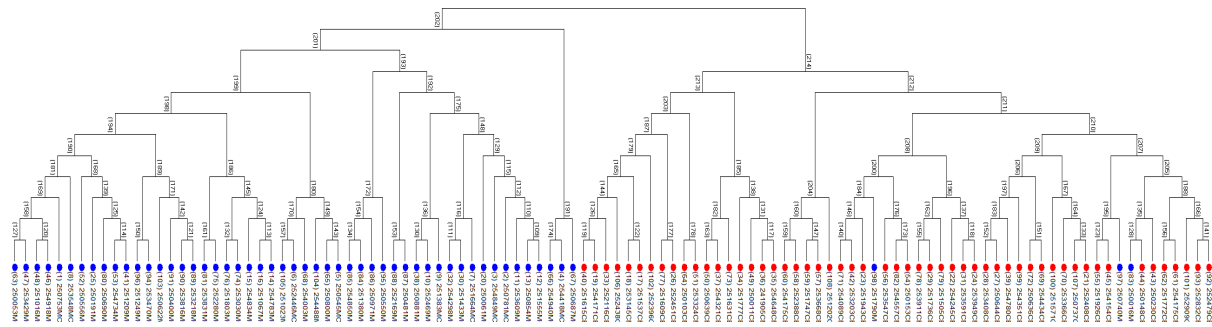


Figure A.1: Structure Analysis of all individuals. Neighbour joining tree for 108 founder individuals in MEGA 8 using POSA (proportion of shared alleles data) from MSA.

A structure analysis was also carried out for 108 individuals to make sure that there were indeed two populations of mice present in the enclosure. As expected, the best result was for K=2, meaning that the genetic analysis confirmed two distinct groups of mice to begin with, and later on the presence of both mixed and pure individuals in all three enclosures (Figure A.2). Mice stayed in enclosures for approximately 10 months and were allowed to establish territories and breed uninterrupted except for monthly measurements of personality and welfare checks. All experiments were run until the second generation of mice, born in the enclosures, had a chance to choose mates (originally called G2 although there were backcrosses).

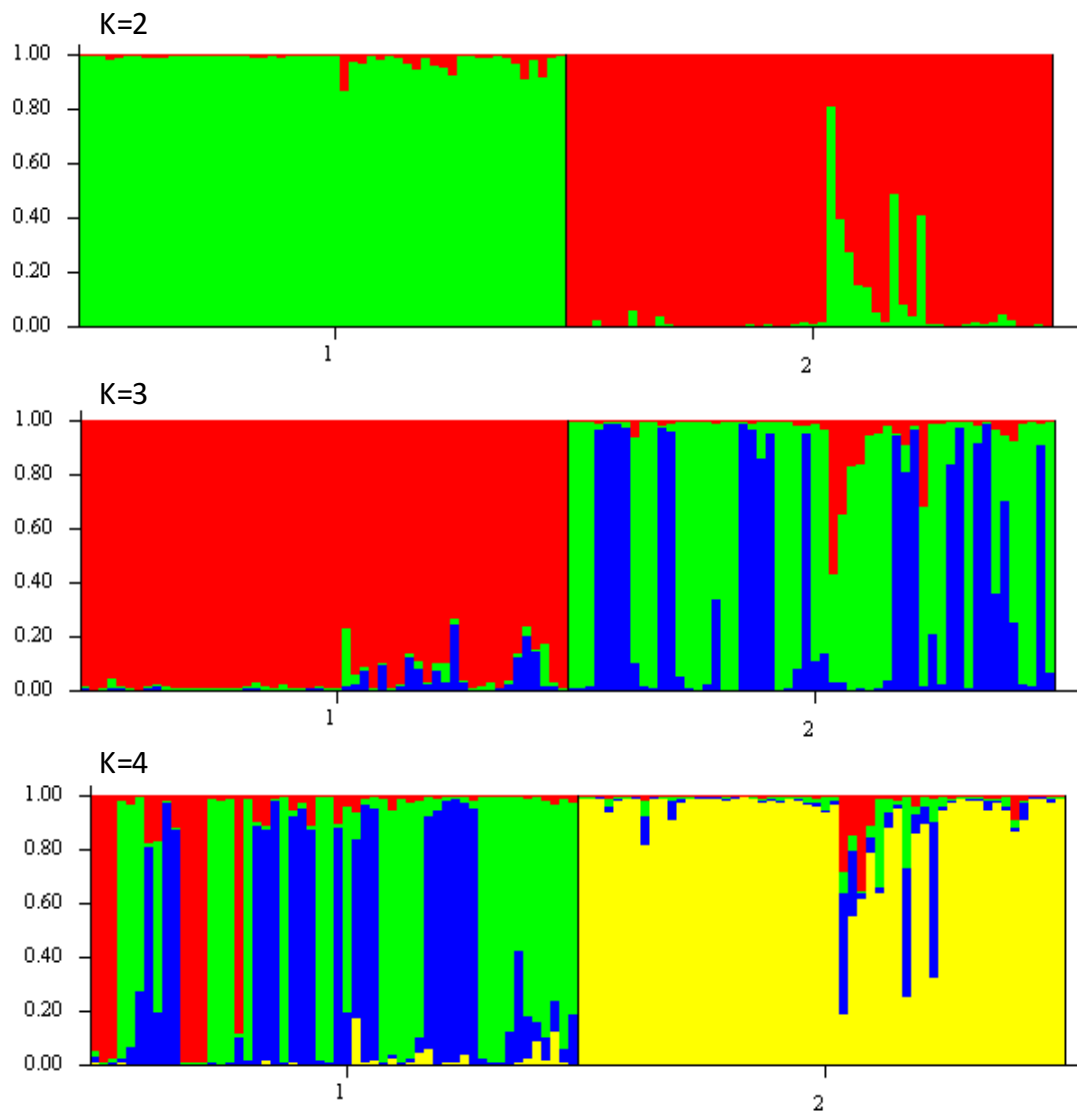


Figure A.2: Structure Analysis of founder individuals. K=2 is the most likely number of clusters underlying the group of mice in the enclosures. These two clusters correspond well with the two populations. Increasing number of clusters would still detect a separation of the two populations, only that one of the two populations would get more structured. Individuals clusters outside the expected populations are the same as the 3 individuals F mice classified as G in Figure A.1. Due to uncertainty of the origin of these three individuals, they were excluded from the behavioural analysis.

The population density for each experiment varied from enclosure to enclosure over time. Generally, the populations grew all the time, more detailed information about the development of the populations is shown in Table A.1 below. The sex ratio for each enclosure started at 1:1. At the end of the experiment, the sex ratio shifted towards a higher frequency of females in enclosure A 83:60 (58% female, X-squared = 3.6993, df=1, p-value = 0.05444), enclosure B, 98:60 (61% female, 8.6101, df=1, p-value = 0.003) and enclosure C, 82:64 (56% female, X-squared = 2.2192, df=1, p-value = 0.1363). The final population

density in the enclosures was less than that of a comparable previous experiment (Montero *et al.* 2013), where after an experiment running 6.5 months (compared to our 10), population density was between 11.2 and 12.9 mice per m² (compared to our between 7.7. and 8.4).

Table A.1: Development of the population in each enclosure over time. Enclosure A (room 5) was (17.6 m²) and enclosure B (room 13) and C (room 14) were the same size (19 m²) and enclosure C was slightly smaller (17.6 m²). All enclosures started with an even number of MC and CB mice. Density is measured in adult mice per square meter. Total is the number of adult mice counted in the enclosure during a monitoring (Mon) except for the final, which includes pups, living at the time. Mon= Monitoring, Gen1 = 1st generation offspring produced by founder individuals, Gen 2= 2nd generation offspring produced by the first generation of mice born in enclosures or backcrosses. Please note that totals are not cumulative and represent the count during the final monitoring period.

Enclosure	Mon	Female	Male	Founder	Gen 1	Gen 2	Density	Total
A	Start	18	18	36	-	-	2.0	36
	1	18	18	36	-	-	2.0	36
	2	18	18	36	-	-	2.0	36
	3	17	18	35	-	-	2.0	35
	4	22	22	34	10	-	2.5	44
	5	38	29	34	33	-	3.7	67
	6	44	40	29	37	18	4.8	84
	7	48	48	27	37	32	5.5	96
	8	43	44	18	31	38	4.9	87
	Final	83	60				8.1	143
<hr/>								
Enclosure	Mon	Female	Male	Founder	Gen 1	Gen 2	Density	Total
B	Start	18	18	36	-	-	1.9	36
	1	17	18	35	-	-	1.8	35
	2	17	18	35	-	-	1.8	35
	3	17	17	34	-	-	1.8	34
	4	23	17	32	8	-	2.1	40
	5	26	24	32	18	-	2.6	50
	6	45	34	32	16	31	4.2	79
	7	40	44	29	17	38	4.4	84
	8	56	38	14	12	68	4.9	94
	Final	98	60				8.4	158
<hr/>								
Enclosure	Mon	Female	Male	Founder	Gen 1	Gen 2	Density	Total
C	Start	18	18	36	-	-	1.9	36
	1	17	17	34	-	-	1.8	34
	2	17	15	32	-	-	1.7	32
	3	27	16	32	11	-	2.3	43
	4	39	31	30	40	-	3.7	70
	5	49	35	25	59	-	4.4	84
	6	56	39	25	62	8	5.0	95
	7	58	46	22	59	23	5.5	104
	8	63	67	13	35	82	6.8	130
	Final	82	64				7.7	148

Parents were assigned to individuals that survived until the age of 3 weeks (approximately a weight of 12 grams). This means that fetal mice and pups which did not reach adulthood were not sampled. 344 offspring out of 452 total were assigned unequivocally to certain parents. This left 66 individuals (13%) which could not be assigned to parents using Cervus 3.0. The following table (Table A.2) shows the count of individuals in each population in the founder Generation (G0) , first generation born in enclosures (G1), second generation born in enclosures (G2), third generation born in enclosures (G3) and backcross combinations. Unassigned individuals were dropped from future analysis.

Table A.2: Number of individuals per generation in each enclosure. The actual generation and count of all individuals that reached adulthood during the experiment in all three enclosures is shown. Furthermore, based on their parentage assignment, the actual generation and cross type for every mouse is included below. *Total assigned includes G0 mice

Generation	Enclosure A	Enclosure B	Enclosure C	Total
G0	36	36	36	108
G1	47	8	31	86
G2	13	18	19	50
G3	7	21	6	34
Backcross G0/G1	19	20	18	57
Backcross G1/G2	15	31	40	86
Backcross G0/G2	5	13	13	31
Total Assigned	142	147	163	452
Unassigned	23	33	10	66
All Mice	165	180	173	518*

At the end of the experiment, the overall count of individuals from each population background varied between enclosures (Figure A.3).

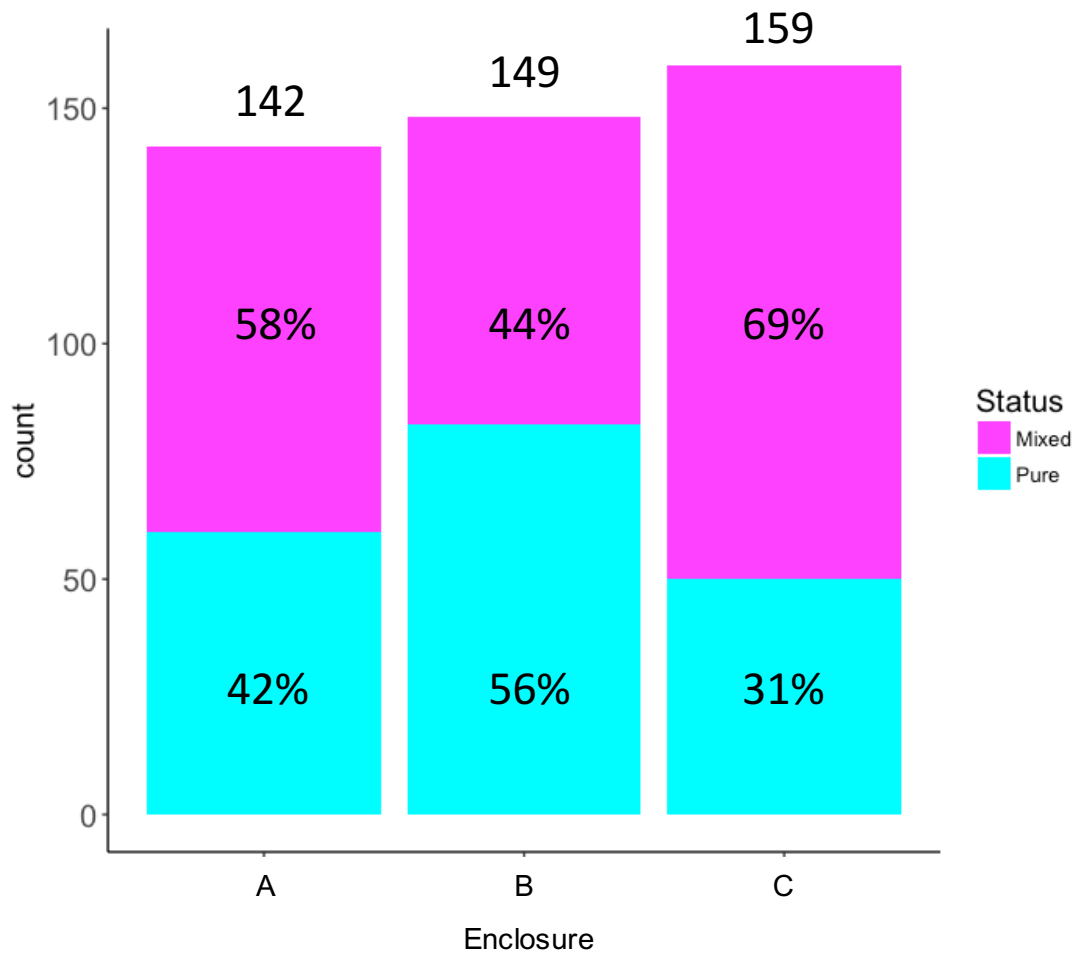


Figure A.3: Pure vs mixed individuals by enclosure. 450 individuals including founder mice and mice with assigned identity for all three enclosure experiments.

B. Mate Choice

Structure Offspring

A Structure analysis for all samples including the parents of the founder generation was performed using the same criteria as for the founder samples alone. Again, K=2 was identified as the most likely number of clusters (Figure B.1).

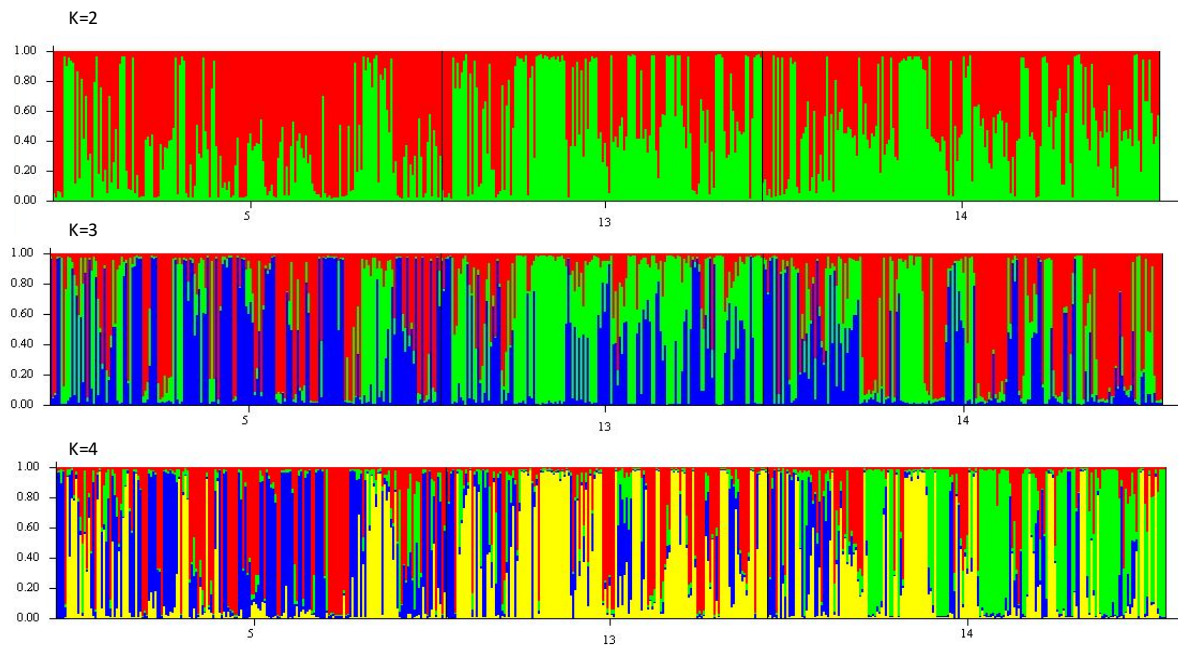


Figure B.1: Structure Analysis of all individuals. 518 individuals from all three enclosures (A, B and C). For comparison, K=2-4 is shown, where K=2 is the most likely number of clusters.

To track mate choice in the enclosures, we used parentage analysis to identify the number of individuals per generation (see section population development). Further we used this information to identify possible mate choice strategies (e.g. assortative mating, paternal or maternal matching) of the parents.

Founder individuals' mate choice and number of first generation offspring

51 founder to founder (G0 to G0) successful mating events were counted overall. This count is not based on litters, rather successful mating events themselves, to include multiple paternity. This means for females one litter may consist of multiple successful mating events if there are multiple fathers. For males, this means if they produce multiple offspring within the same litter, it is only counted once. Successful mating events between founder and founder parents were found in all monitoring events, including the final one. 27 mating events were in enclosure A, 8 from enclosure B and 16 from enclosure C. Enclosure B was very different from the other two enclosures, as it was subject the more disturbances (repairs) at the beginning of the experiment. It is possible that this may have delayed breeding. All enclosures produced 27 mixed status mating events (FG or GF) and 24 pure (FF or GG) showing that overall there was no tendency toward assortative/disassortative mating for the founder generation (combined: chi-squared= 0.17, df=1, p-value=0.67). Mating events between founder individuals are shown in by enclosure figure B.2. To determine whether or not there was assortative mating by

enclosure, pure vs mixed status pairings were compared by enclosure. For enclosure A (chi-squared= 0.33, df=1, p-value=0.56), B (chi-squared= 0.5, df=1, p-value=0.47) and C (chi-squared= 0.25, df=1, p-value=0.61), separately, there was no detectable mating pattern present. Comparing these results to a previous experiment (Montero, 2010), these results are similar to the patterns found in experiments I and IV, whereas they found a disassortive mating pattern in exp III and assortative pattern in exp II. This suggests that the results from the founder to founder mating events can be pooled. No analysis was done for t-haplotype, since it was not present in the parent population.

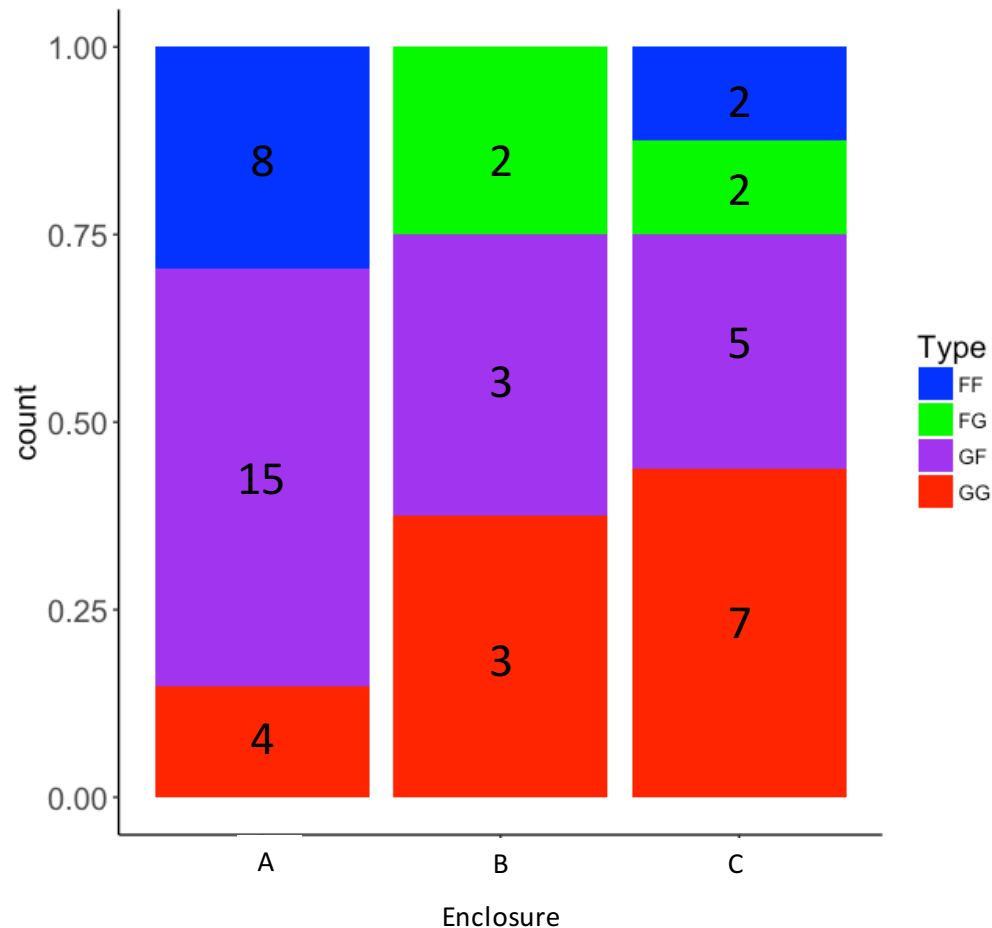


Figure B.2: Proportion of successful mating events by enclosure. There were 51 total founder to founder mating events, 27 from enclosure A, 8 from enclosure B and 16 from enclosure C.

Focusing on the G1 generation to other individuals of the G1 generation, there were 40 successful mating events in all 3 enclosures combined (compared to 156 successful mating events in Montero, 2010). Figure B.3 shows the successful mating counts for G1 to G1 mating events. The patterns can be compared to Montero *et al.* (2013). Montero *et al.* (2013) compared the results from observed successful G1 to G1 mating events to results generated from a null model for random mating based on the genotypes which were present in the enclosures at the time. Since we had so few successful mating events, it was not possible to replicate this statistic exactly. Instead, a null model for random mating based on potential mates was calculated based on the potential parents present at each time point (monitoring periods), and the average of the proportion was taken for assortative matches (GG to GG, FF to FF, FG to FG and GF to GF), maternal matches (GG to GF, FF to FG, GF to GG and FG to FF), paternal matches (GG to FG, FF to GF, FG to GG and GF to FF) and disassortative (GG to FF, FF to GG, GF to FG and FG to GF). When the overall proportion of observed successful mating events for each of the four categories was compared to the expected proportion of successful mating events, the null hypothesis could not be rejected (chi-squared=12, df=9, p-value=0.21).

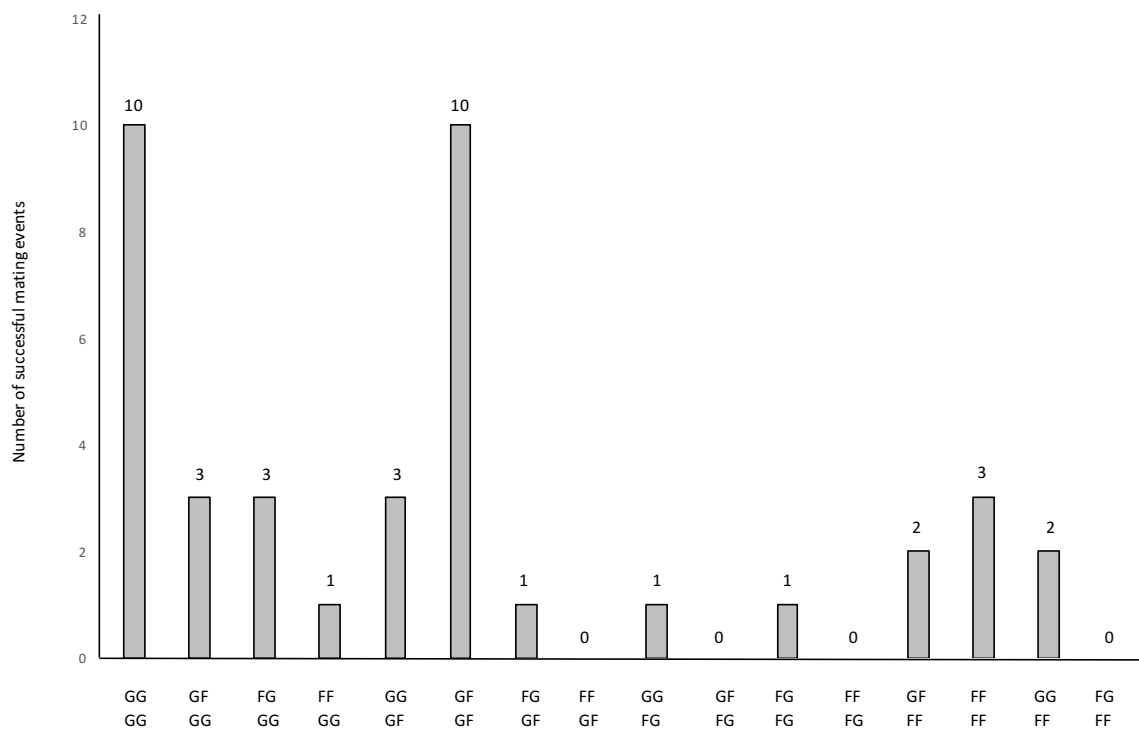


Figure B.3: Number of successful of mating events for G1 individuals. The Y axis shows the count of mating events, the top two letters on the X axis show the genotype of the mother, and the bottom two letters show the genotype of the father.

Montero found that mate choice in her experiments was heavily influenced by the father in hybrids. This means that if an individual had a father from the F population, they would choose to mate with an individual who also had a father from the F population (or were pure MC). There were clearly more GF/GF and GG/GG successful mating events than other types with an observed proportion of 7% of all successful mating events for GF/GF and 3% for GG/GG observed versus expected proportions of 0.5% GF/GF and 2% GG/GG expected.

In order to analyse backcrosses between the G1 and G0 mice, the same strategy Montero, 2010 used was implemented. Namely, female G1 mating with G0 males were considered separately to G0 females mating with G1 males. There were 20 cases of G1 females mating events with G0 males (Figure B.4.A) and 29 mating events between G1 males and G0 females (Figure B.4.B). For female G1 mice mating with G0 mice, there was a slight pattern towards paternal matching (chi-squared=7.57, df=3, p-value=0.06). There was a similar pattern for male G1 mating with female G0 (chi-squared=8.12, df=3, p-value=0.043). This analysis cannot be used to infer paternal/maternal mating; however, since results are likely confounded by inbreeding.

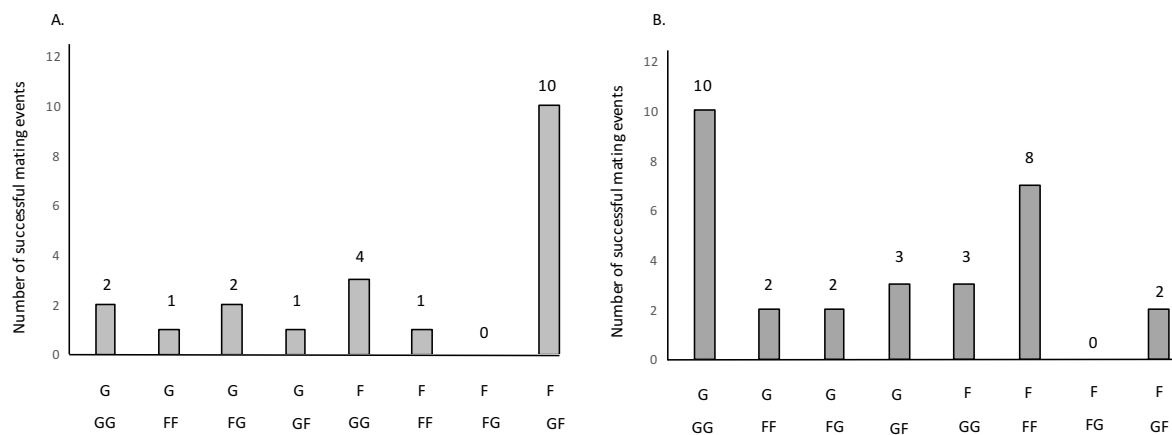


Figure B.4: Backcrosses between G1 and G0 mice. A. Backcrosses between female G1 mice to male G0 mice (20 events total) and B. Backcrosses between male G1 and female G0 mice (29 total events).

Reproductive Success

113 females out of 280 (40%) and 115 (48%) out of 238 males produced offspring that survived to be adults (more than 12 grams) during the enclosure experiment. Individuals that did not survive until

adulthood were not recorded; therefore reproductive success is based only on offspring which survived to the age where they could reproduce themselves.

Reproductive success by population

Focusing first on the founder mice and counting all, 27 females produced offspring (out of 54). Of these, 16 were CB/G females and 11 were F. There was no difference between female F and CB/G mice (chi-squared=1.38, df=1, p-value=0.23). 28 unique G0 fathers produced offspring. Of these, 13 were CB/G and 15 were F (chi-squared=0.142, df=1, p-value=0.70).

For G1 female mice with a pure status (GG and FF), there were 22 unique mothers. Of them, 15 were pure CB/G and 7 were pure F. There was no significant difference between the two groups, but the p-value was not far from significance (chi-squared=2.90 df=1, p-value=0.08). Montero, 2010 found that there was a significant difference between GG and FF females, where more GG females produced offspring than FF females. Comparing mixed background mice (GF and FG), there was a total of 21 unique mothers, 5 were FG and 16 were GF. There were significantly more GF mice than FG (chi-squared=5.76 df=1, p-value=0.01). This result was different from Montero, 2010, where no difference was found between GF and FG females. Finally, female G1 mice were grouped together into the categories of pure vs. mixed female G1 mice (22 FF and GG) and (21 GF and FG). Similar to Montero, 2010, there was no significant difference between in reproductive success between pure and mixed G1 females (chi-squared=0.02, df=1, p-value=0.87).

In the case of G1 males, there were 27 unique fathers, 11 of which were F (FF) and 16 of which were CB/G (GG). There was no significant difference between these two groups (chi-squared=0.92, df=1, p-value=0.33). 18 mixed (FG and GF) G1 males produced offspring. Of these, 6 were FG and 12 were GF. There was no significant difference between these two groups either (chi-squared=2, df=1, p-value=0.15). Comparing pure vs. mixed background mice directly, there was still no significant difference between pure and mixed G1 males (chi-squared=1.8, df=1, p-value=0.18). Similar to Montero 2010, all groups of G1 males appeared to have the same chance of producing offspring as the others.

Individual reproductive success

Similar to Montero, 2010, individual reproductive success was compared by measuring the number of offspring an individual produced (absolute and relative), the number of mating events (absolute and relative), for females the number of offspring per litter and for males the number of offspring per mating event. Relative scores are scaled by the number of days an individual was in the experiment, to correct for time spent in the experiment. For example, individuals which were not in the experiment as long have less opportunity to produce offspring. As above, mice raised in cages (G0) and mice raised in the experiment (G1) were analysed separately.

First, focusing on G0 males which produced offspring (Figure B.5), there was no significant difference in the number of offspring produced (Figure B.5.A, Wilcoxon Test, $n=27$, $p\text{-value}=0.92$) or the relative number of offspring produced (Wilcoxon Test, $n=27$, $p\text{-value}=0.54$) between the two populations (Figure B.5.B). There was also no significant difference between the populations when it came to number of mating (Figure B.5.A, Wilcoxon Test, $n=27$, $p\text{-value}=0.41$) or the relative number of mating (Figure B.5.B, Wilcoxon Test, $n=27$, $p\text{-value}=0.92$). Finally, there was no difference in the overall number of offspring produced per litter by population (Figure B.5.C, Wilcoxon Test, $n=27$, $p\text{-value}=0.43$). These results were similar to those reported in Montero, 2010.

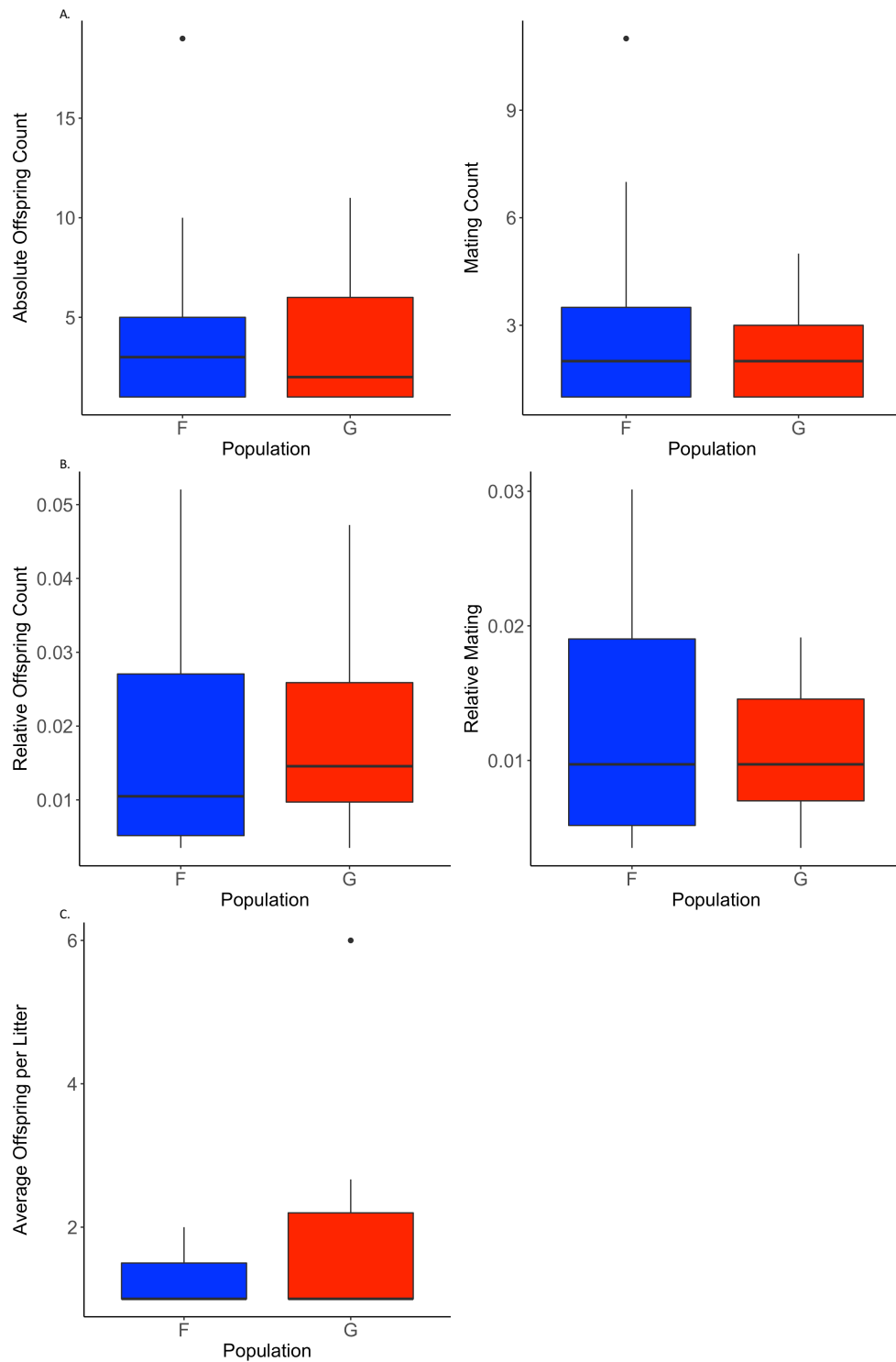


Figure B.5: Reproductive success of G0 males by population. A. the absolute number of offspring (right) and the number of successful mating events (left) by G0 male mice from F (F) and CB/G (G) populations. B. the relative number of offspring produced (right) and the relative mating count (left) for G0 male mice from F and CB/G G populations and C. the average number of offspring per litter for G0 male mice by population F and G.

Second, looking at the results for female G0 mice (Figure B.6), there was no significant difference between the absolute number of offspring produced by female G0 mice (Figure B.6.A, Wilcoxon test, $n=27$, $p\text{-value}=0.18$) or the relative number of offspring produced (Figure B.6.B, Wilcoxon test, $n=27$, $p\text{-value}=0.27$). Nor was there a difference between the absolute number of successful mating events by population (Figure B.6.A, Wilcoxon test, $n=27$, $p\text{-value}=0.40$) or relative number of successful mating events (Figure B.6.B, Wilcoxon test, $n=27$, $p\text{-value}=0.51$). Finally, there was no difference in the average number of offspring produced per litter by F or G female G0 mice (Figure B.6.C, Wilcoxon test, $n=27$, $p\text{-value}=0.20$). These results were also consistent with Montero, 2010.

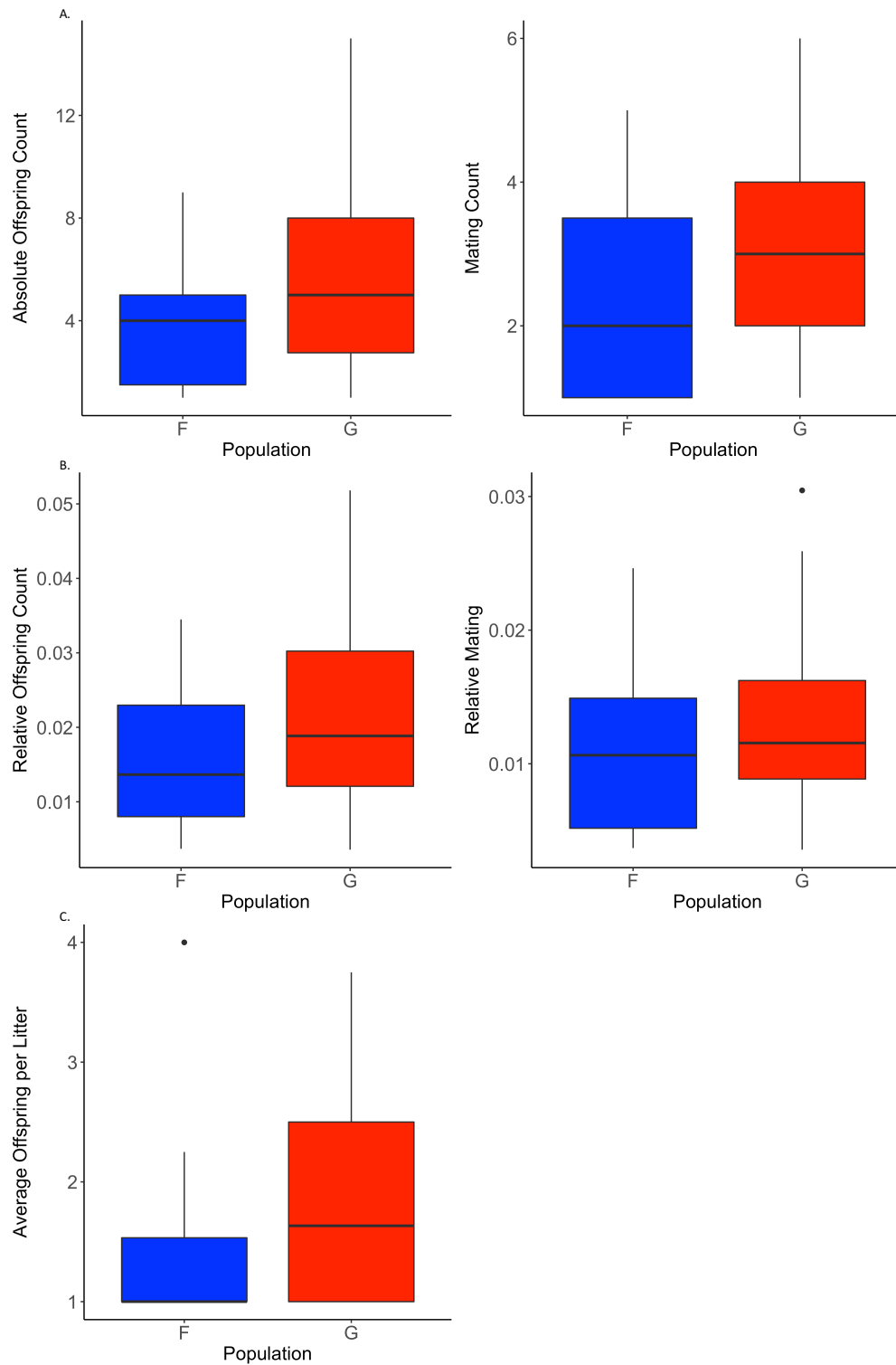


Figure B.6: Reproductive success of G0 females by population. A. the absolute number of offspring (right) and the number of successful mating events (left) by G0 female mice from F (F) and CB/G (G) populations. B. the relative number of offspring produced (right) and the relative mating count (left) for G0 female mice from F and CB/G G populations and C. the average number of offspring per litter for G0 female mice by population F and G.

Next, considering the mice born in the enclosures, the absolute and relative number of offspring for G1 males and females was considered, as well as the number of mating and the relative number of mating events and finally the average number of offspring per litter. This was done first for all four types of offspring (FF, GG, FG and GF) and later for individuals of pure (FF and GG) and mixed (GF and FG) status.

Starting with G1 males (Figure B.7), there was no significant difference between the types in the absolute number of offspring produced (Figure B.7.A, Kruskal-Wallis chi-squared = 3.6308, df = 3, p-value = 0.30) or the relative number of offspring produced (Figure B.7.B, Kruskal-Wallis chi-squared = 1.7014, df = 3, p-value = 0.63). There was no difference in the mating count (Figure B.7.A, Kruskal-Wallis chi-squared = 3.474, df = 3, p-value = 0.32) or relative mating count (Figure B.7.B, Kruskal-Wallis chi-squared = 0.564, df = 3, p-value = 0.90) and there was no difference in the average number of offspring produced by litter between the four types (Figure B.7.C, Kruskal-Wallis chi-squared = 3.2856, df = 3, p-value = 0.34). These results were also similar to Montero, 2010.

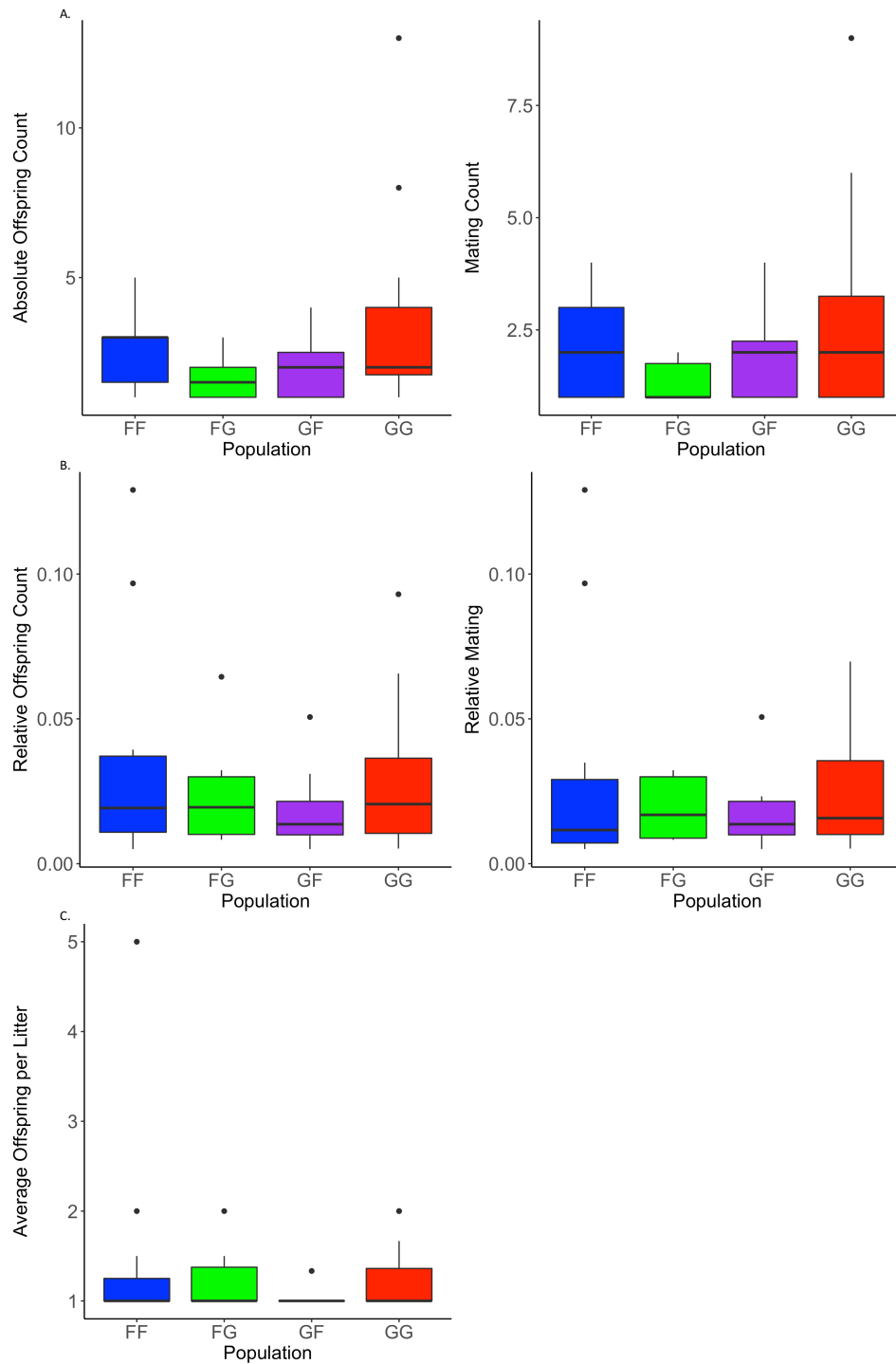


Figure B.7: Reproductive success of G1 males by population background. A. the absolute number of offspring (right) and the number of successful mating events (left) by G1 male mice from pure FF and GG mice, and mixed FG and GF mice B. the relative number of offspring produced (right) and the relative mating count (left) for G1 male mice from pure FF and GG mice, and mixed GF and FG mice C. the average number of offspring per litter for G1 male mice by population FF, GG, FG and GF background.

Next, looking at G1 females (Figure B.8). There was no significant difference between the types concerning offspring count or relative offspring count (Figure B.8.A, Kruskal-Wallis chi-squared = 1.7906,

df = 3, p-value = 0.62 and Figure B.8.B, Kruskal-Wallis chi-squared = 4.0562, df = 3, p-value = 0.25 respectively). Nor was there a difference between the groups in the absolute mating count (Figure B.8.A, Kruskal-Wallis chi-squared = 4.9182, df = 3, p-value = 0.17). There was a significant difference between the relative number of successful mating events between the four types (Figure B.8.B, Kruskal-Wallis chi-squared = 12.183, df = 3, p-value = 0.006). In this case, pure FF females were much more successful than the other types (Figure B.8.B). Finally, there was no significant difference between the average number of offspring per litter by type (Figure B.8.C, Kruskal-Wallis chi-squared = 0.44832, df = 3, p-value = 0.93). These results deviated from those found in Montero, 2010. In the previous study, there was no significant difference between the four types of G1 offspring when focusing on the relative number of mating events; however, there was a significant difference between the absolute number of offspring produced between the four types, but not in the relative values.

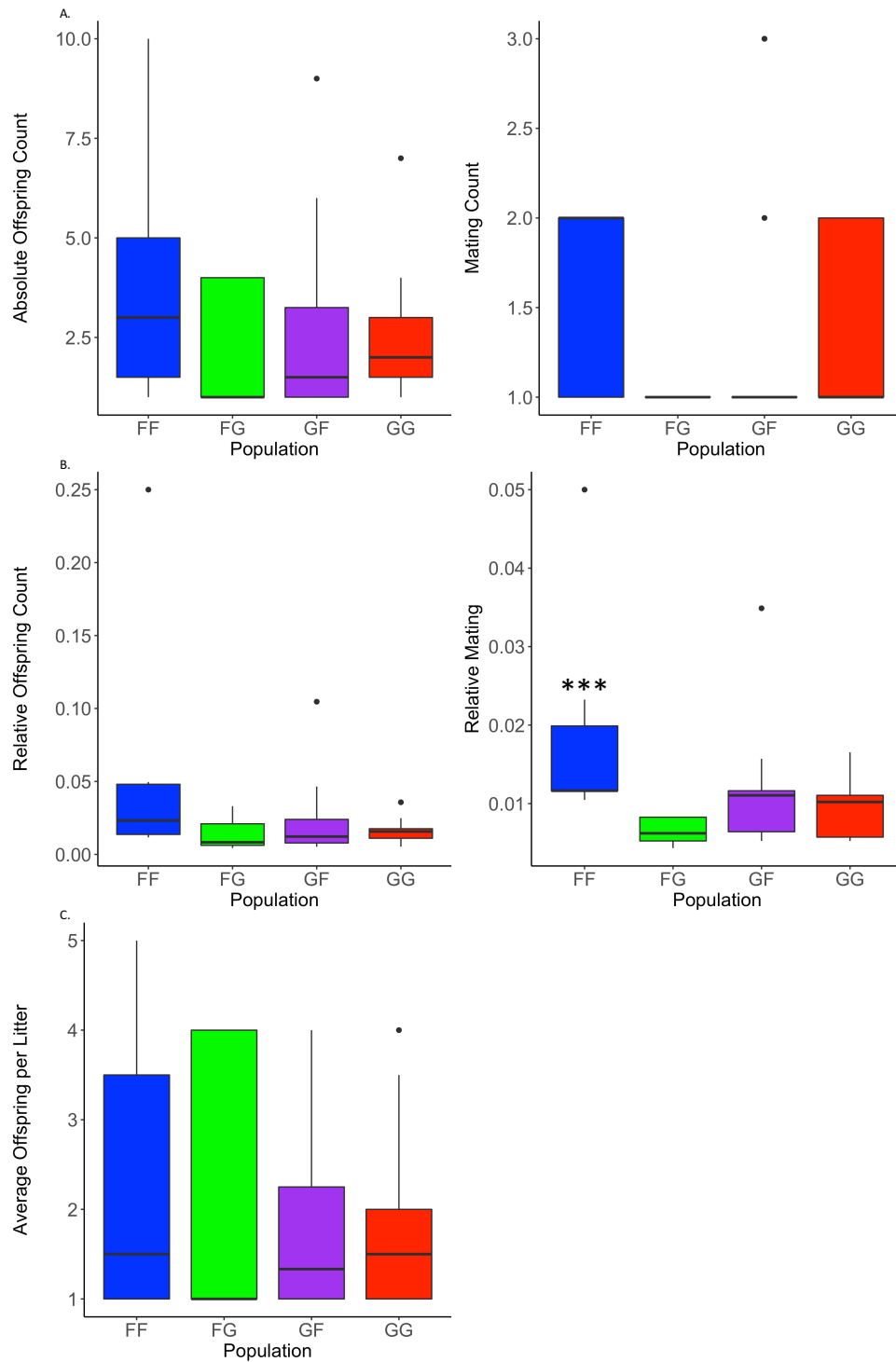


Figure B.8: Reproductive success of G1 females by population background. A. the absolute number of offspring (right) and the number of successful mating events (left) by G1 female mice from pure FF and GG mice, and mixed FG and GF mice B. the relative number of offspring produced (right) and the relative mating count (left) for G1 female mice from pure FF and GG mice, and mixed GF and FG mice C. the average number of offspring per litter for G1 female mice by population FF, GG, FG and GF background.

Now, combining the pure background types (FF and GG) and the mixed types (FG and GF), the same measurements were used to compare G1 mice from pure and mixed status background starting with males and then females. For pure vs mixed status G1 males, the absolute number of offspring produced was approaching significance (Figure B.9.A, Wilcoxon Test, p-value = 0.07), although when it was adjusted for by time in the enclosure, this effect disappeared (Figure B.9.B, Wilcoxon Test, p-value = 0.26). Mating count and relative mating count were also not significantly different (Figure B.9.A, Wilcoxon Test, p-value = 0.21 and Figure B.9.B, Wilcoxon Test, p-value = 0.71 respectively). Finally, the average number of offspring per litter was not significantly different (Figure B.9.C Wilcoxon Test, p-value = 0.20). These results were also similar to Montero, 2010, except that in the previous study, the absolute number of offspring produced was not as close to significance.

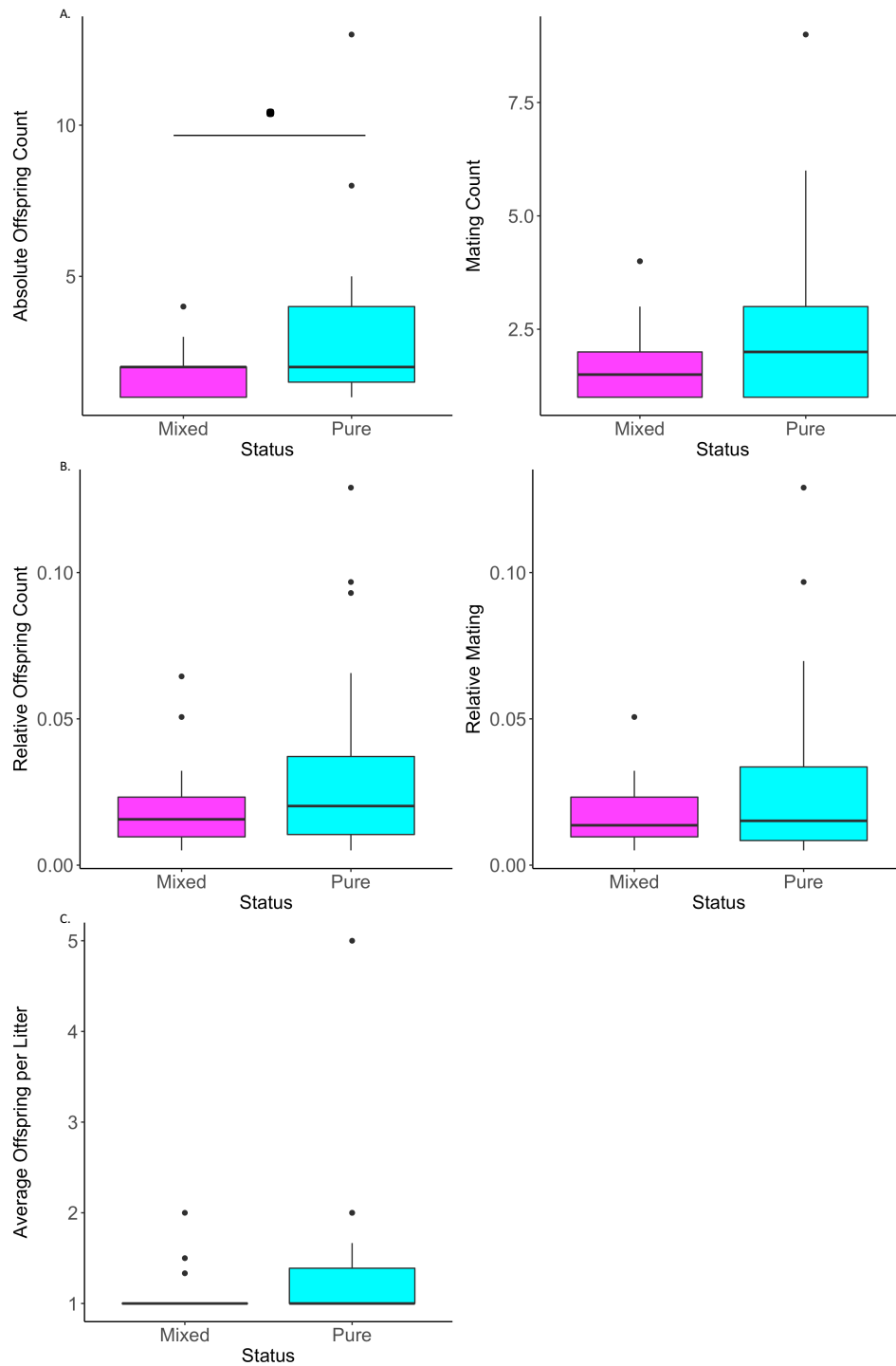


Figure B.9: Reproductive success of G1 males by status. A. the absolute number of offspring (right) and the number of successful mating events (left) by G1 male mice by pure and mixed status mice B. the relative number of offspring produced (right) and the relative mating count (left) for G1 male mice for pure and mixed status mice C. the average number of offspring per litter for G1 male mice by population pure vs. mixed background mice.

Finally, for G1 females by pure vs mixed status (Figure B.10), there was no significant difference for any of the measurements. The absolute number of offspring produced (Figure B.10.A, Wilcoxon Test, p-value = 0.23) and the relative number of offspring produced (Figure B.10.B, Wilcoxon Test, p-value = 0.32) did not differ by status. Mating count was approaching significance Figure B.10.A, Wilcoxon Test, p-value = 0.06), but corrected for the number of days in the experiment (Figure B.10.B, Wilcoxon Test, p-value = 0.24), it was far from significantly different. Finally, the average number of offspring produced per litter was not different between the groups (Figure B.10.C, Wilcoxon Test, p-value = 0.57). Finally, these results were also consistent with Montero, 2010, except that in this case, the mating count was close to significance. The relative mating count however was not.

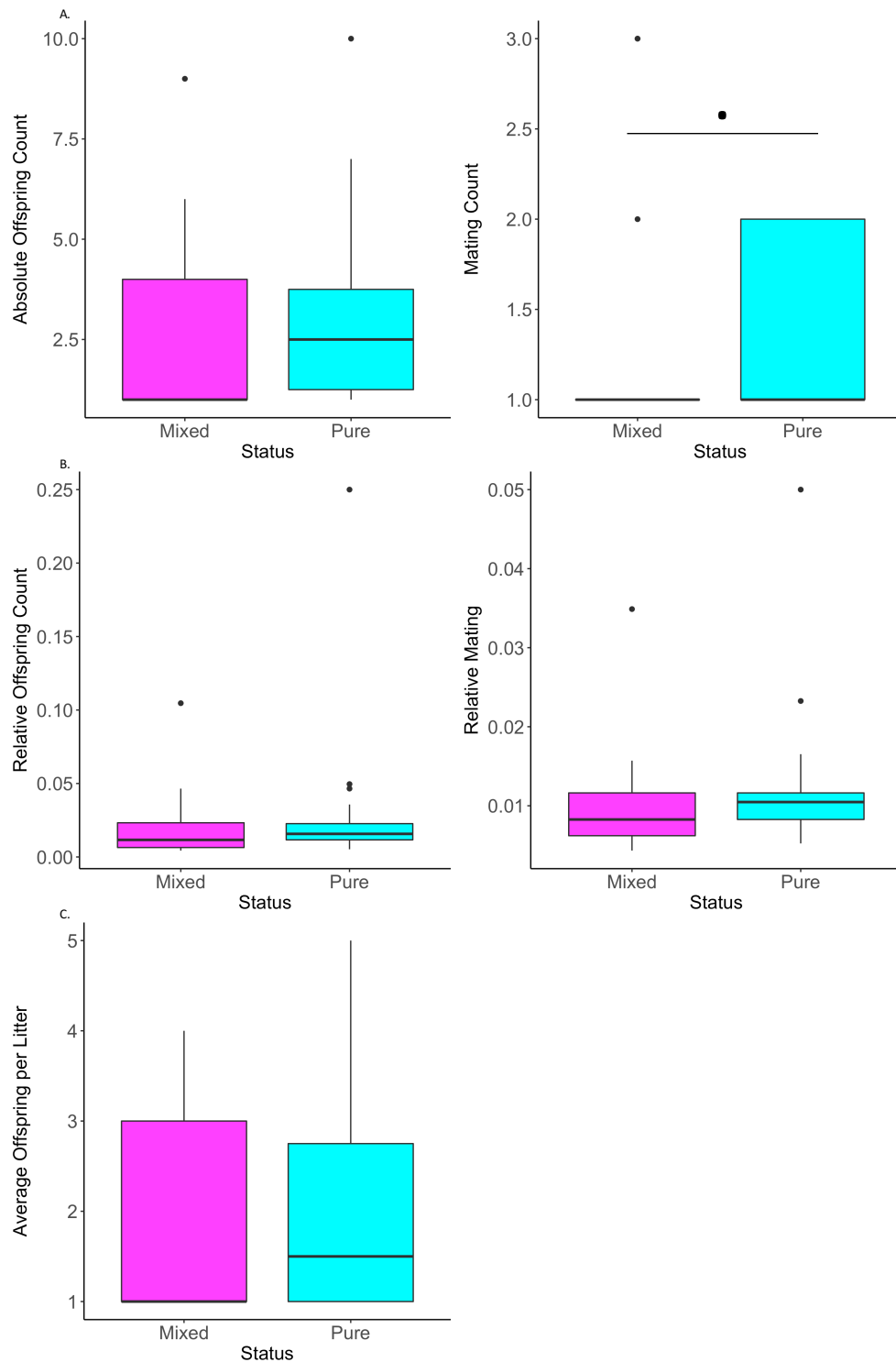


Figure B.10: Reproductive success of G1 females by status. A. the absolute number of offspring (right) and the relative number of offspring produced (left) by G1 female mice by pure and mixed status mice B. the absolute mating count (right) and the relative mating count (left) for G1 female mice for pure and mixed status mice C. the average number of offspring per litter for G1 female mice by population pure vs. mixed background mice.

C. Territory Use

Enclosures were equipped with antennae to study the locations of the mice used during the experiments to track spatial use. This information was recorded to understand how mice use available space and form territories and how they change these over time. It is also interesting to understand the relationship between the scores of first generation of mice born in the enclosures to their parents (founders) and understand which factors influence these scores. Finally, it is also interesting to understand how territory use may influence the number of offspring produced by mice during the experiment, by calculating mean scores of territory-use and comparing it to the absolute number of offspring produced by mice during the experiment.

House Occupancy

Throughout the experiment monitoring events were carried out to assess the health and welfare of the mice in the experiment and carry out personality tests (section B.3). Each house was removed from the enclosure and checked. During these monitoring events records were kept about the locations of the mice and house occupation. House occupancy and the location of mice were then compared to the information assessed by the antennae. The results from enclosure 14 show that all shelters were being visited throughout the experiment (Figure C.1.A), even though during monitoring some houses were found to be empty (Figure C.1.B). The same is true for enclosures 13 and 5, details are shown in the supplemental information Figure S1 and Figure S2.

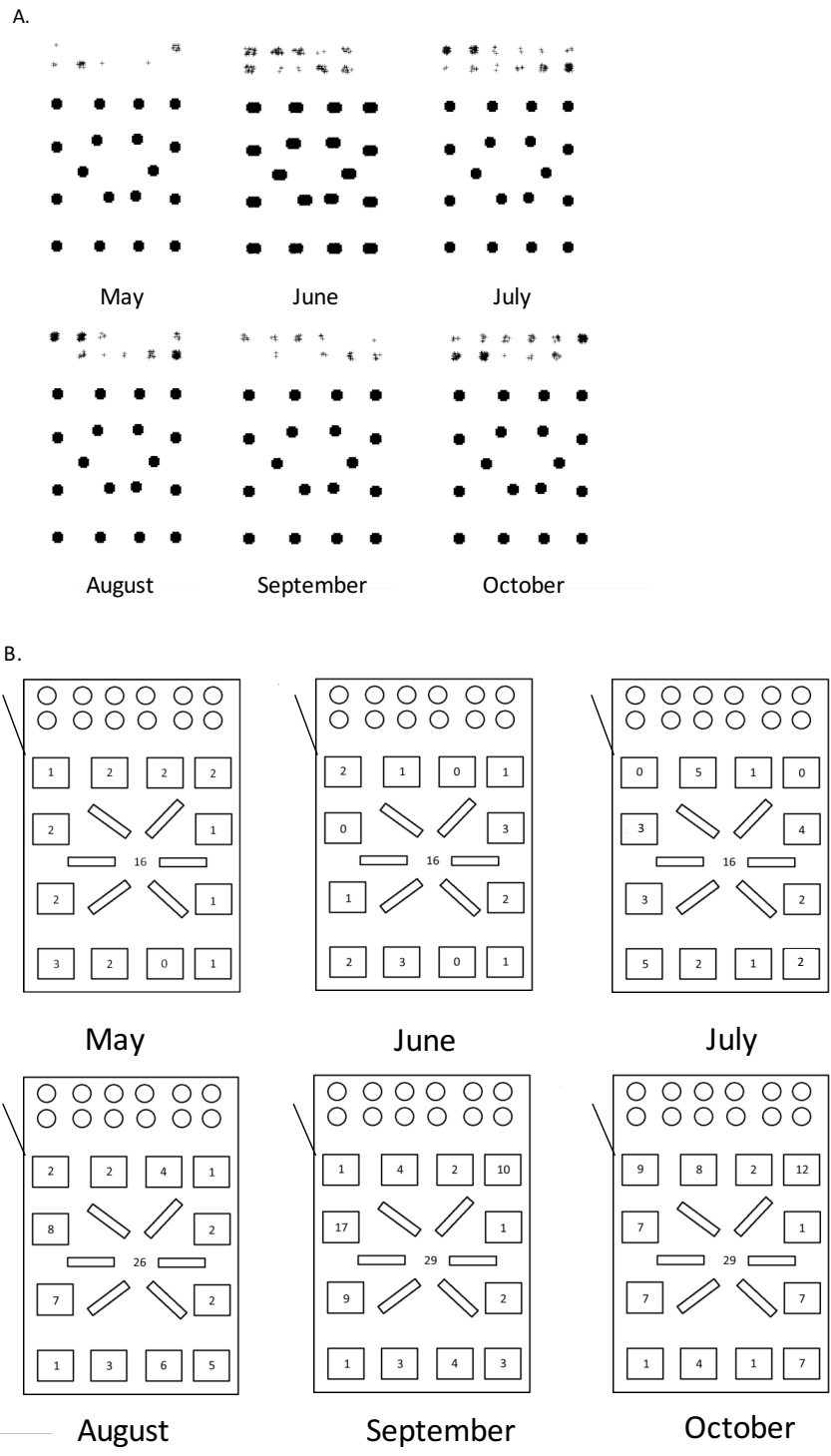


Figure C.1: Locations used by mice. A. Activity of all individuals, as sum of the reads per antennae per month, is given for enclosure C. All potential shelters (included water bottles, houses and pipes) were visited throughout the course of the experiment. To visualize data, some “jitter” was added to create small variances in the data points so that the antennae reads do not appear all in the same position. B. The locations mice were found during monitoring events over the course of a year in enclosure C. The number in the center of the graph shows the number of individuals which were caught outside of the houses.

Further, antenna read maps were made for all 353 individuals which were in contact with antennae (not all mice registered on antennae). The number of antenna reads per month was counted and used as a proxy for arena activity (AA). Minimum convex polygons (MCPs) were estimated for where the mouse spent 95% of their time (providing they had at least 5 location changes over the course of the month) to estimate the area of the home range size (HRS). Antenna read maps show the locations of individuals not caught in houses during monitoring events can also be ascertained (Figure C.2.A). These plots were also used to compare the mice caught in houses during the monitoring periods to the locations in which they spent the most of their time. These results show that the mice were often caught within the estimated home range, but not always (Figure C.2.B). Maps for all individuals by month are included in the supplemental information S3. This also shows that some of the mice that were caught in traps appear to have very stable territories indeed and may have simply been outside at the time we completed the monitoring.

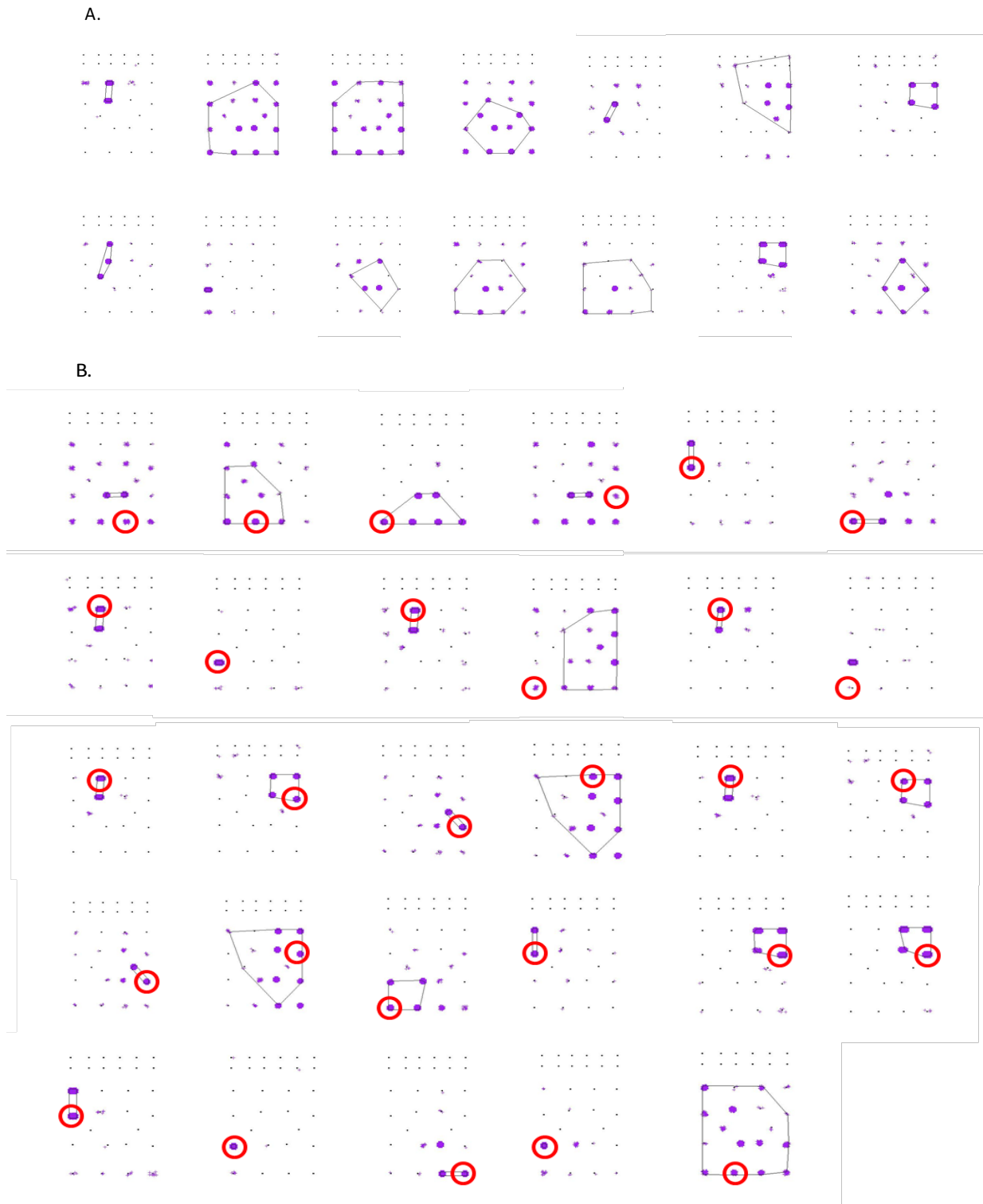


Figure C.2: Example of HRS maps. A. Maps of the spatial use of individuals which were caught in traps during July (classified as free range mice) in enclosure C. Purple markings indicate antenna hits and polygons enclose the area in which the mouse spent 95% of its time. A minimum of 5 antenna reads were necessary to make polygons and therefore no plots were generated for two mice. B. The houses in which mice were caught during monitorings were circled in red. Purple markings indicate antenna hits and polygons enclose the area in which the mouse spent 95% of its time.

Territory Measurements

HRS and AA were used to study the territorial behaviour of mice. In addition, site selectivity (SI) was also measured. This measurement was derived from the standard deviation derived from the proportion of antennae used out of a possible 30 locations, divided by the maximum possible standard deviation, in this case 0.1826 (see Material and Methods for details). This means that there is a maximum score of 1, if a mouse spends 100% of its time in one location and a minimum score of 0, if the mouse spends an equal amount of time at all locations. Correlation plots were made to explore the relationship between these measurements by individual. Focusing on founder generation mice, there was an inverse relationship between the scores for HRS and AA (Figure C.3.A, $n=188$, Spearman rank Correlation, $\rho = -0.17$, $p\text{-value} = 0.015$). There was also an inverse relationship between the SI score and the HRS (Figure C.3.B, $n=188$, Spearman rank correlation, $\rho = -0.58$, $p\text{-value} < 0.001$) Finally, Figure C.3.C shows no significant relationship between the AA scores and the SI measurements for individuals over time (C.3.C, $n=188$, spearman rank correlation, $\rho=0.178$, $p\text{-value}=0.016$), although there was a positive trend between the two measurements.

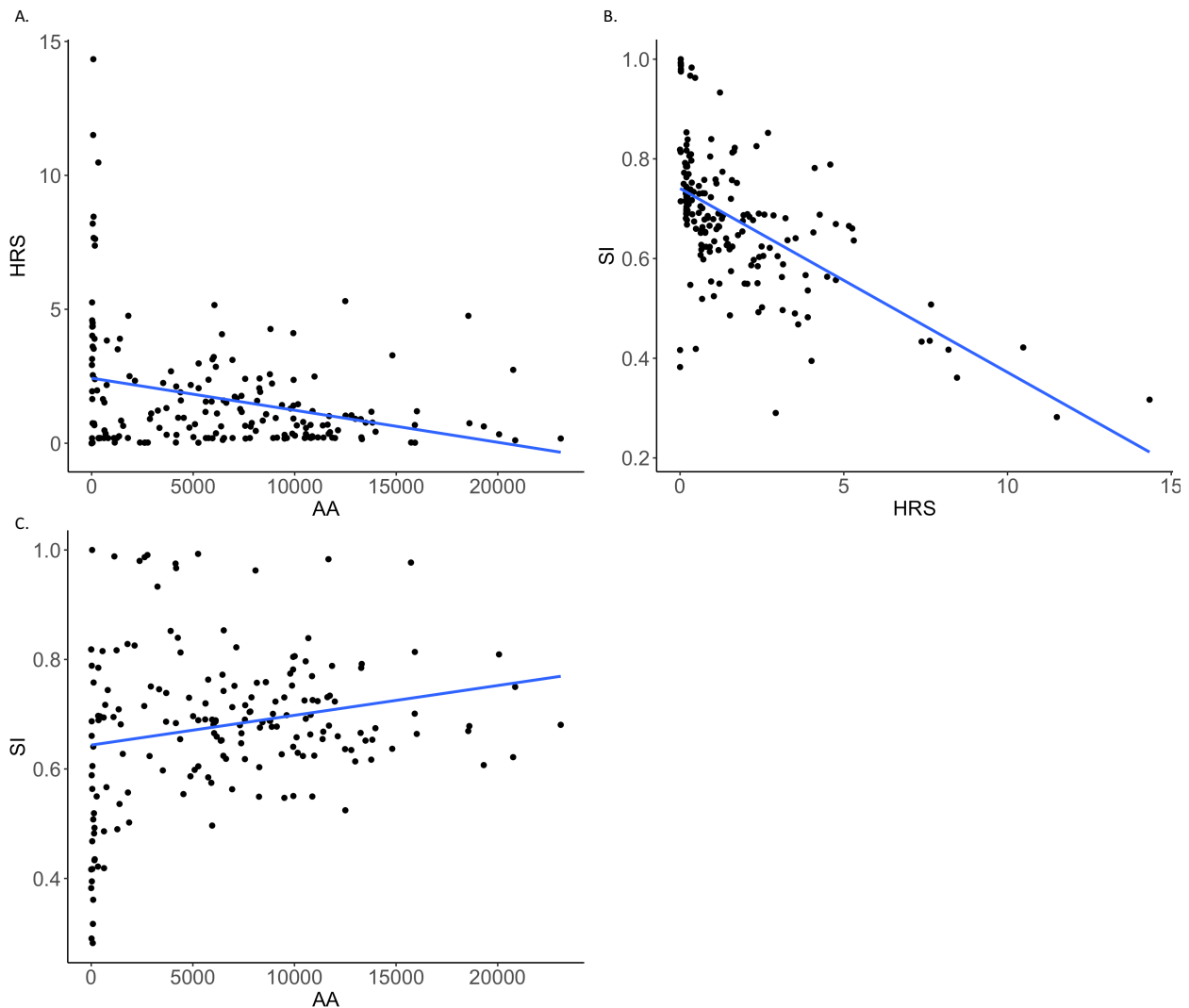


Figure C.3: Relationship between territory scores. A. Average AA derived from antenna reads versus HRS (the territory size) of individuals. B. SI compared to HRS and C. SI compared to AA for n=188 mice.

Mean Territorial Scores over Time

The following figures show the mean scores for AA and HRS by month and SI by monitoring between for sex, status (pure or mixed) and enclosure. Figure C.4 shows the overall changes in AA over time, Figure C.5 shows HRS and Figure C.6 shows the overall changes in SI by monitoring period. The first mixed mice appear in the enclosures after 1.5 month, therefore there are no mixed mice during the first-time. It is also important to note that during the first two months of the experiment, enclosure 13 suffered from power outages and was only repaired in July.

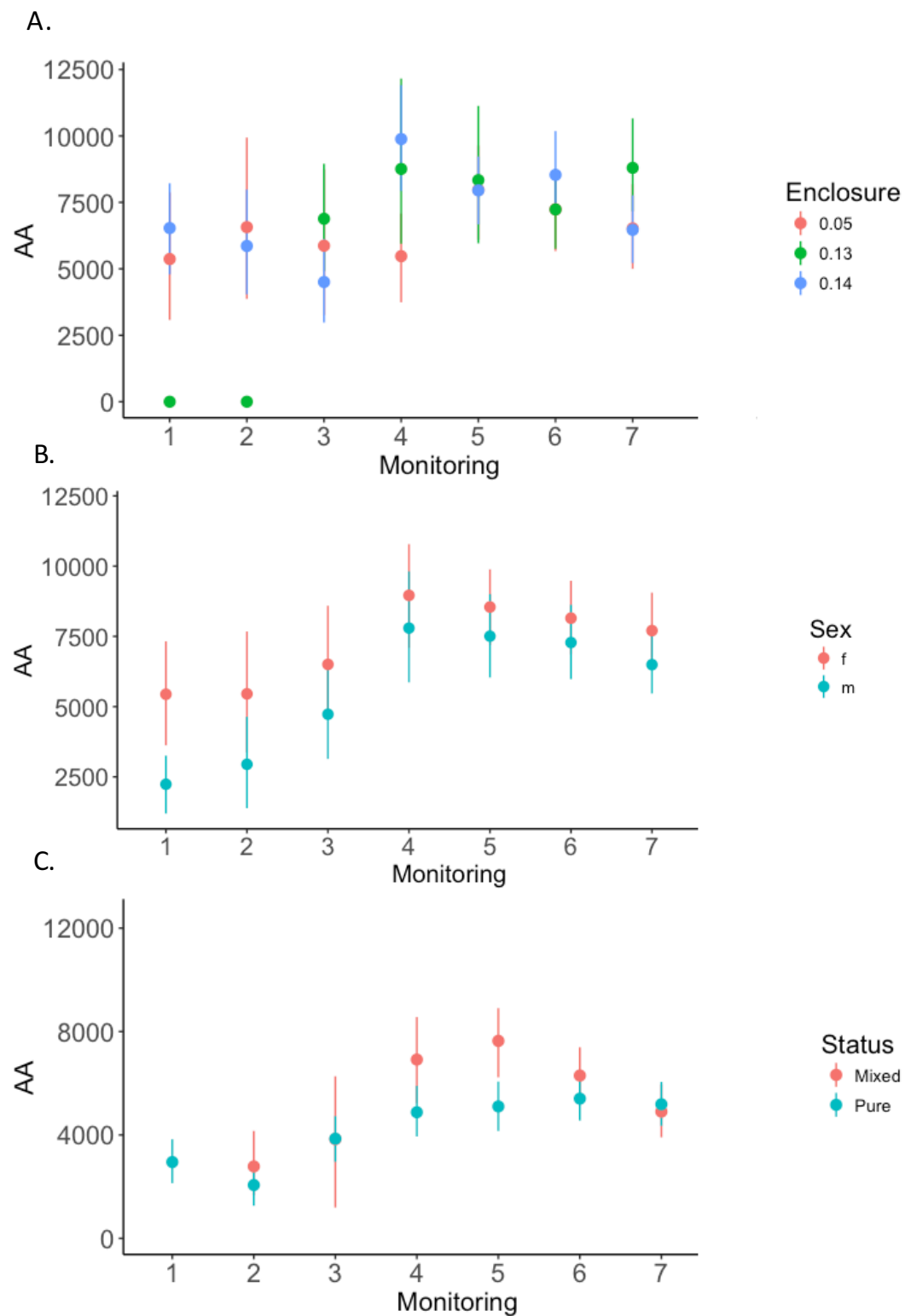


Figure C.4: Mean AA scores for 7 monitoring events. A. AA over time by Enclosure B. Sex and C. Status (mixed or pure lineage). These scores were calculated based on monitoring (1-7) and include scores from 188 mice.

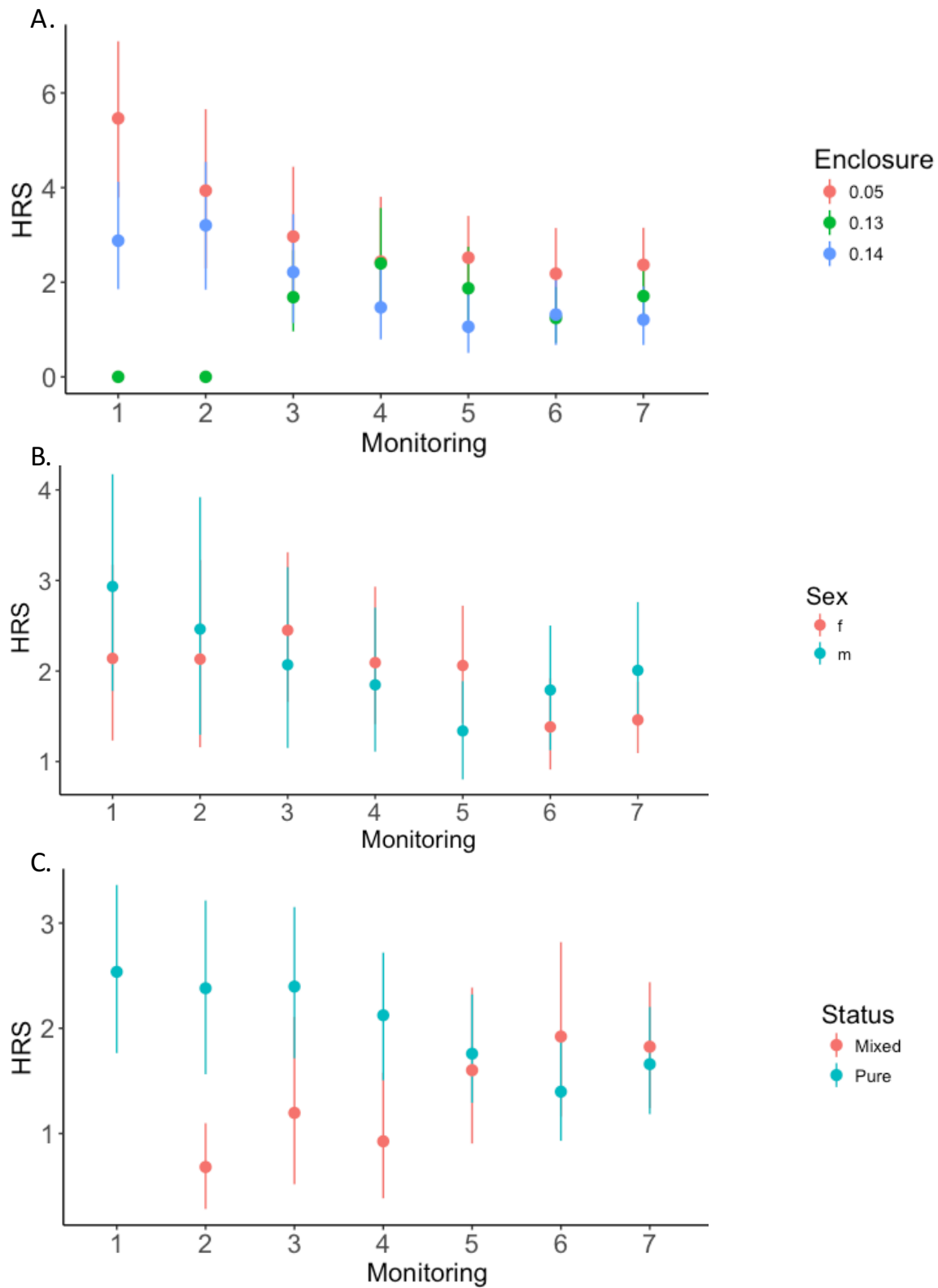


Figure C.5: Mean HRS scores for 7 monitoring events A. HRS over time by Enclosure B. Sex and C. Status (mixed or pure lineage). These scores were calculated based on monitoring (1-7) and include scores from 188 mice.

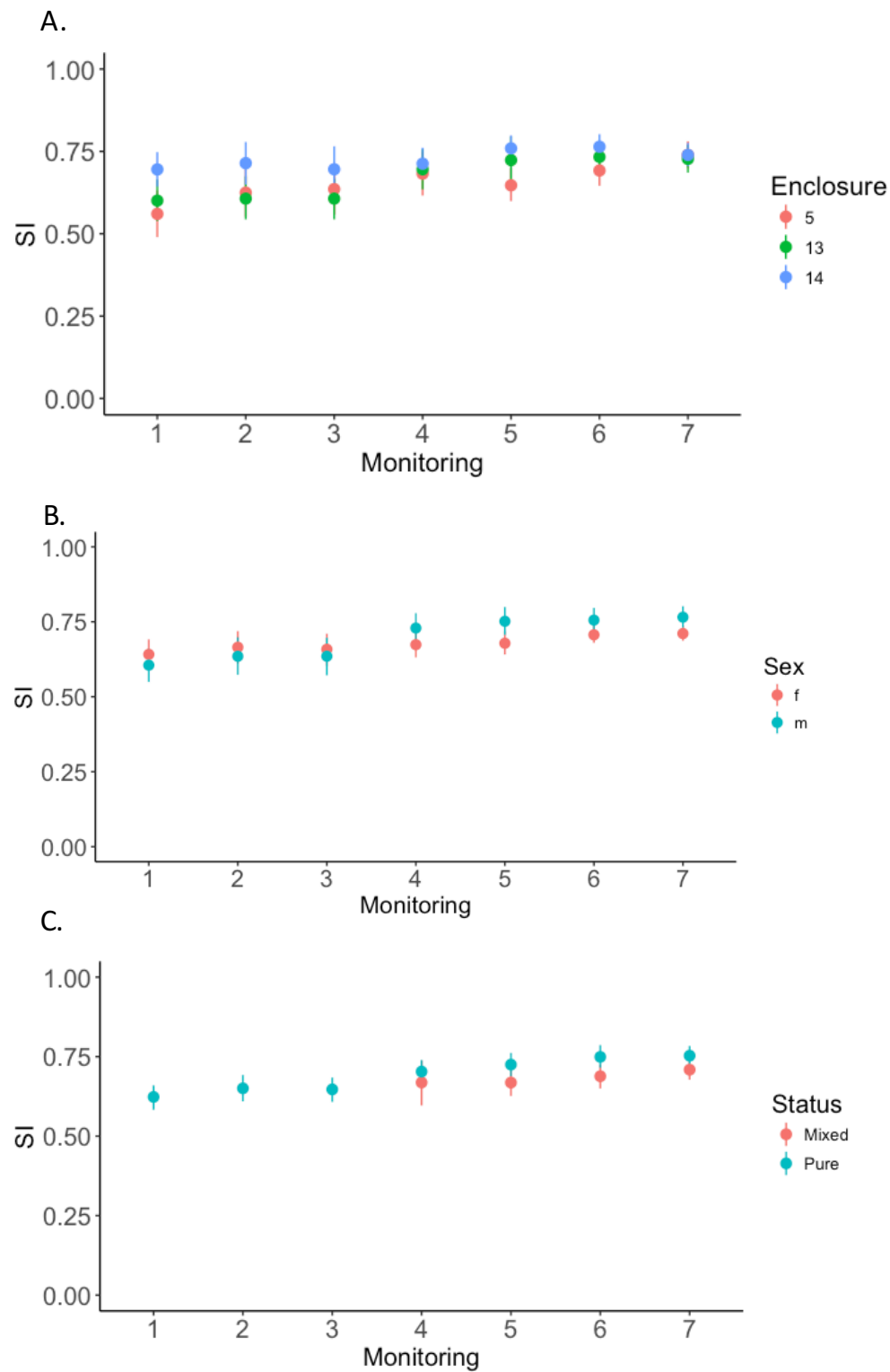


Figure C.6: Mean SI scores for 7 monitoring events A. SI over time by Enclosure B. Sex and C. Status (mixed or pure lineage). These scores were calculated based on monitoring (1-7) and include scores from 188 mice.

Relationship between Parental and Offspring Territory Use

To study the relationship between the offspring and parental mice, the territorial scores of founder mice were compared to those of the first generation of mice, born in the experiment. Mean scores for AA, HRS and SI for the entire year were calculated to estimate the average territory use for offspring and parents. Spearman correlations were carried out to determine the relationship between the scores the mother, the father and the mean score of the parents to the offspring. For AA, there was no relationship ($\rho=0.19$, $p\text{-value}=0.1226$), nor was there a relationship between the offspring's and father's AA score (Figure C.7 B. $n=67$, Spearman rank correlation, $\rho=0.24$, $p\text{-value}=0.1436$). There was a trend towards a positive relationship between the mean AA scores of the parents and the offspring AA score (Figure C.7 C. Spearman rank correlation, $\rho=0.22$, $p\text{-value}=0.079$).

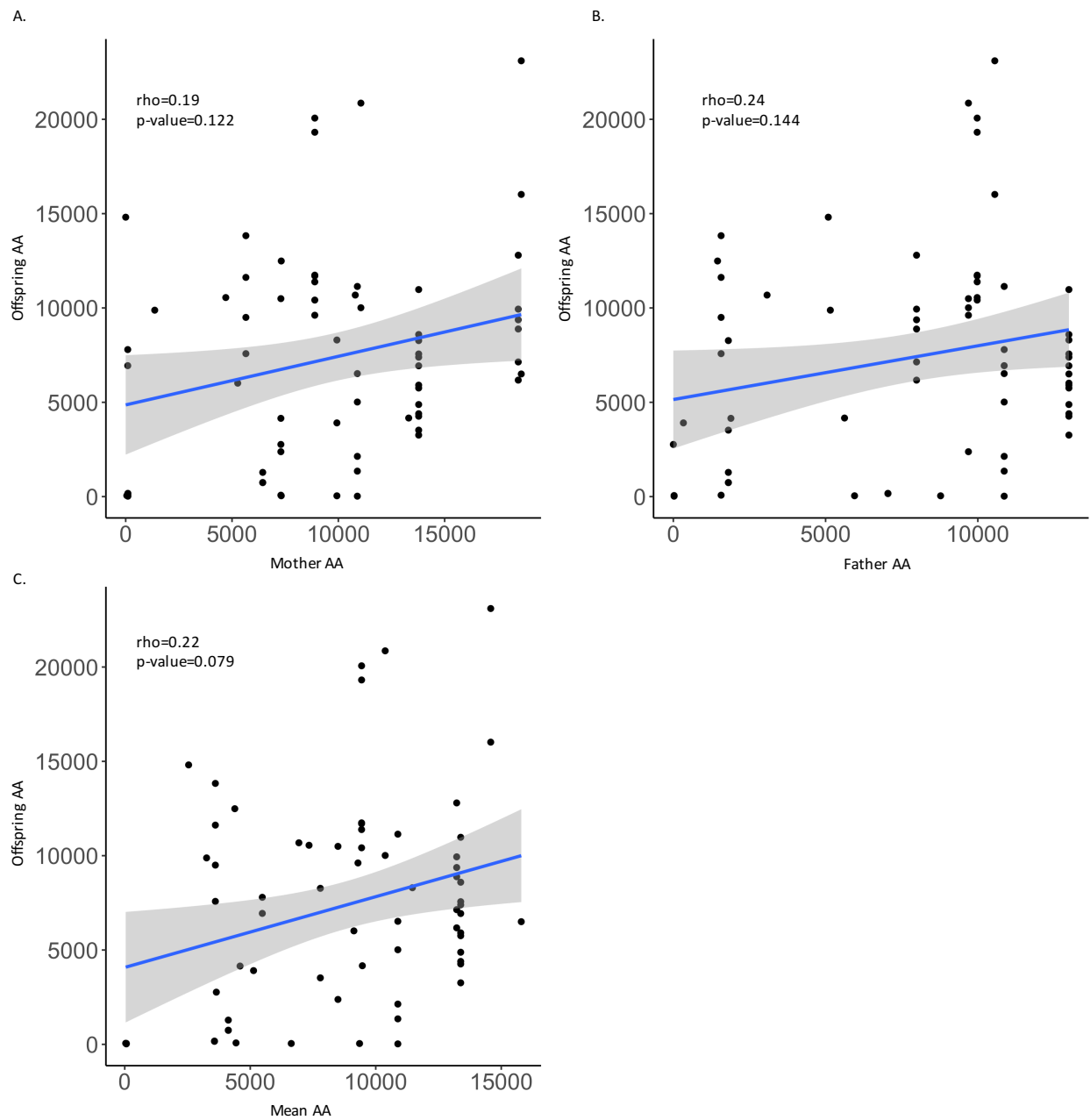


Figure C.7: G1 AA scores compared to parents. The mean AA scores for the first-generation of mice born in the experiment (Offspring AA) scores compared to A. the mother's AA score, B. the father's AA score and C. the mean AA score for both parents for 67 mice.

There was a trend toward a relationship between the mean HRS score of the offspring and the mean score for the mother (Figure C.8.A. spearman rank correlation, $\rho=0.22$, $p\text{-value}=0.069$) although the relationship between the mean HRS score for the offspring and father were strongly correlated (Figure C.8.B. spearman rank correlation, $\rho=0.47$, $p\text{-value}<0.001$). Finally, the mean scores for offspring HRS

and the mean HRS of both parents were also correlated (Figure C.8.C. spearman rank correlation, $\rho=0.37$, $p\text{-value}=0.002$).

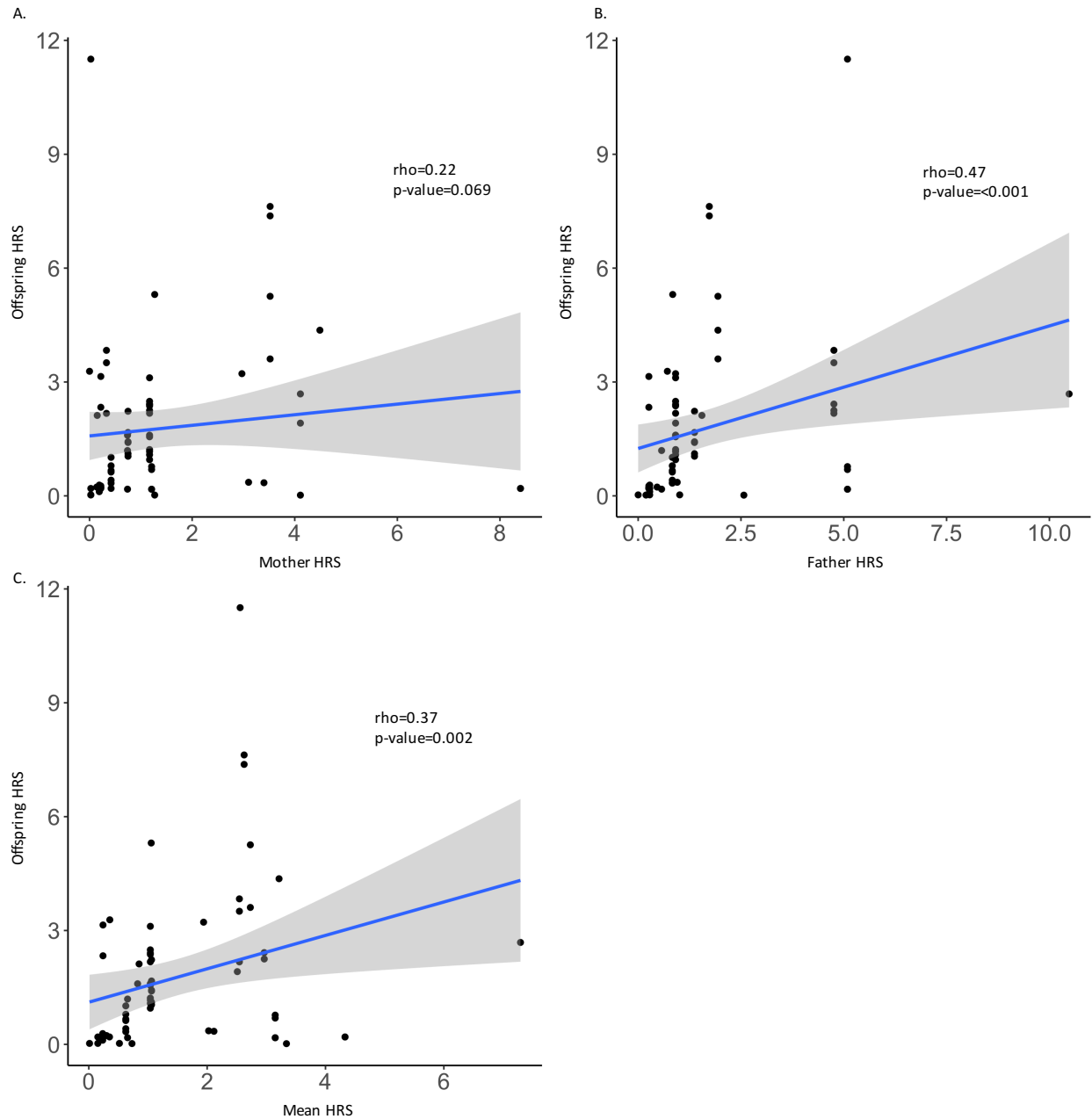


Figure C.8: G1 HRS scores compared to parents. The mean HRS scores for the first-generation of mice born in the experiment (Offspring HRS) scores compared to A. the mother's HRS score, B. the father's HRS score and C. the mean HRS score for both parents for 67 mice.

The mean SI score of first generation offspring born in enclosures was compared to the mean scores of their Mother, Father and the mean score for bother parents combined. In this case, there was a correlation between the mean offspring scores and that of the mother (Figure C.9.A. spearman rank correlation, $\rho=0.292$, $p\text{-value}=0.016$) but there were no relationships between the mean offspring SI and that of the father (Figure C.9.B. spearman rank correlation, $\rho=0.08$, $p\text{-value}=0.49$) or the mean SI score for both parents (Figure C.9.C. spearman rank correlation, $\rho=0.058$, $p\text{-value}=0.64$).

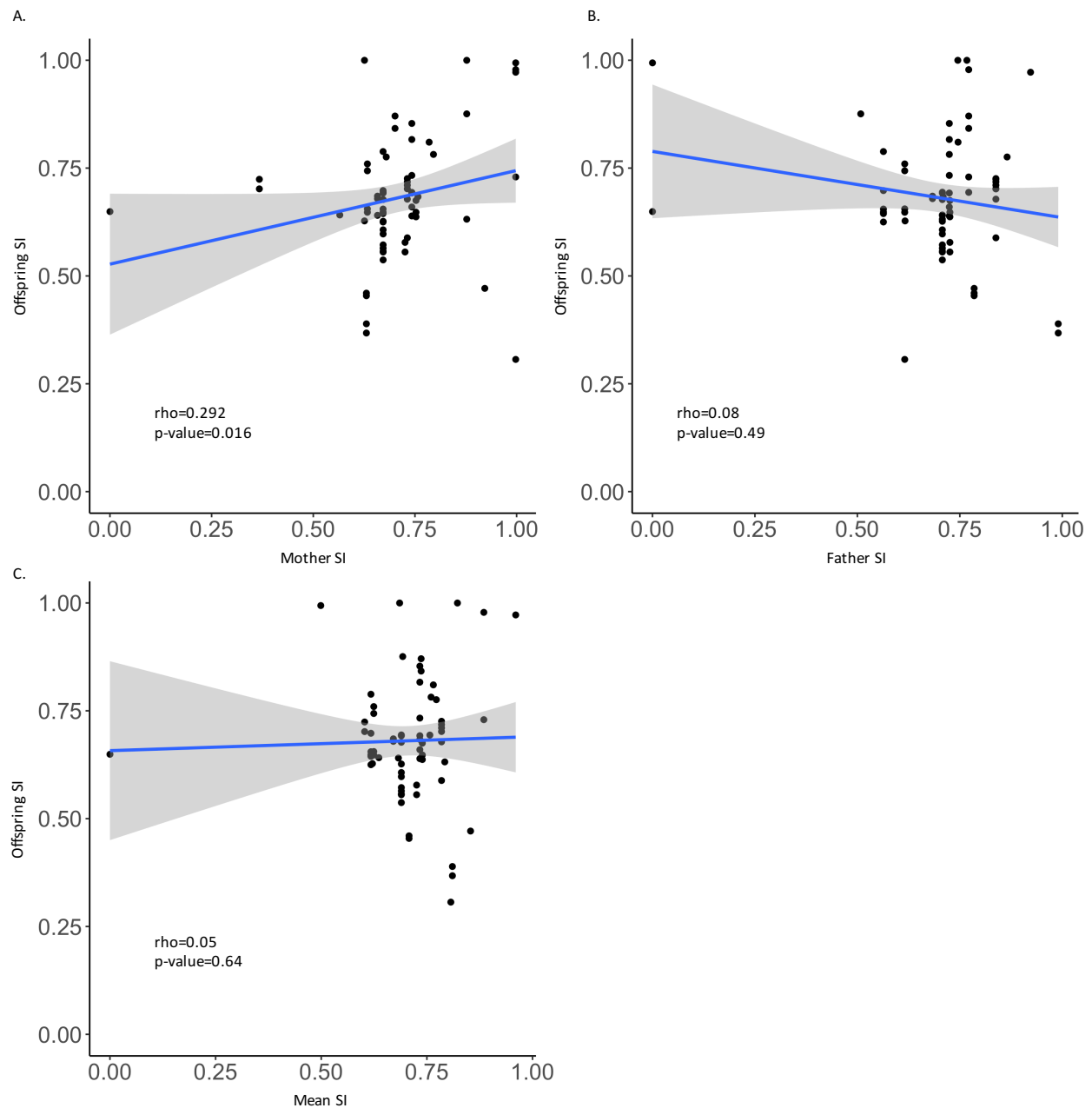


Figure C.9: G1 SI scores compared to parents. The mean SI scores for the first-generation of mice born in the experiment (Offspring SI) scores compared to A. the mother's SI score, B. the father's SI score and C. the mean SI score for both parents for 67 mice.

To further investigate differences in territorial behaviour based on sex, population background and status (mixed vs pure), population of mother and population of father were included as potential influencing factors (Table C.1). By using generalized mixed effects models sex, status and the population origins of the parents were found to influence the AA score. Pure mice had higher scores than mixed

mice overall, regardless of sex (Figure C.10) and there were differences in AA overall between male and female mice (Figure C.10.A). Female mice had higher AA than male mice. Mice with a CB/G Mother (C.10.B) and a CB/G father (C.10.C) had higher AA than mice with F parents.

Table C.1: Generalized mixed effects model for AA scores. AA scores of offspring were compared by fixed factors of sex, status (pure vs. mixed) and the population origin of their mother and father.

Fixed Effect	Estimate	SE	Z value	P-value (Wald)
Intercept	5.723	0.866	6.60	<0.001
Sex	-0.015	0.004	-3.14	<0.001
Status	3.116	0.296	10.51	<0.001
Population Mother	-1.224	0.283	-4.32	<0.001
Population Father	3.215	1.112	2.89	0.003
Sex: Status	-0.3154	0.009	-32.96	<0.001
Random Effects				
Father ID	5.96	2.44		
Mother ID	1.28	3.57		

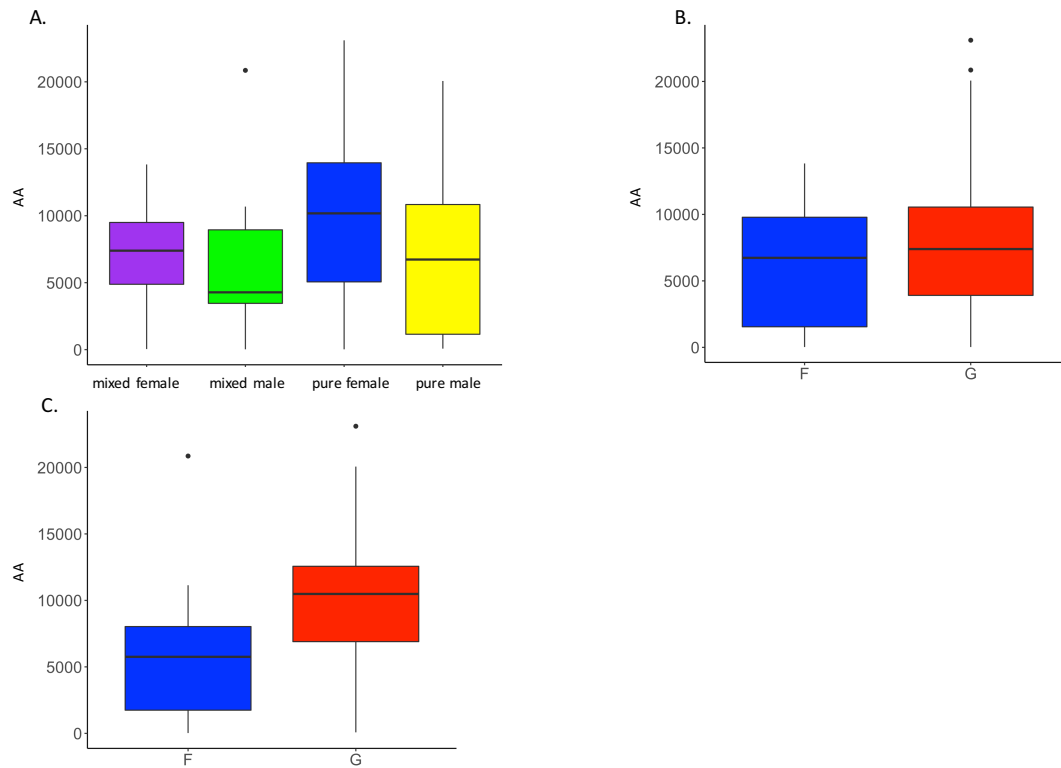


Figure C.10: A. The interaction between AA scores of first generation offspring based on their status (pure vs mixed) by sex B. The AA scores of first generation offspring based on the population origin of their mother F (MC/F), G (CB/G) C. The AA scores of first generation offspring based on the population origin of their father F (MC/F), G (CB/G) for 67 mice.

A similar model was used to find the factors influencing the HRS score (Table C.2). The simplest model to explain relationships between HRS scores of offspring and fixed effects showed that HRS was influenced by status (pure vs mixed) (Figure C.11.A) as well as the population of their mother (Figure C.11.B). In this case, there was no influence of sex on the HRS score, but there were differences between pure and mixed background mice, where mixed individuals had higher HRS than pure ones. Scores were also influenced by the population origin of the mother, but not the father. Individuals with a CB/G mother had higher HRS scores than those with MC/F mothers.

Table C.2. A linear mixed effects model for square root transformed HRS scores of offspring to their parents. In the simplest model, there was an effect of status and the population of the mother on the territory size of offspring. Mother was dropped as a random effect (variance was 0). The random effect of enclosure and father influenced to score.

Fixed Effect	Estimate	SE	t value	P-value
Intercept	0.576	0.866	6.60	<0.001
Status	3.116	0.296	10.51	<0.001
Population Mother	-1.224	0.283	-4.32	<0.001
Random Effects			R	
Father ID	0.576	0.759	0.92	
Enclosure	0.069	0.263	0.59	
Residual	0.048	0.221		

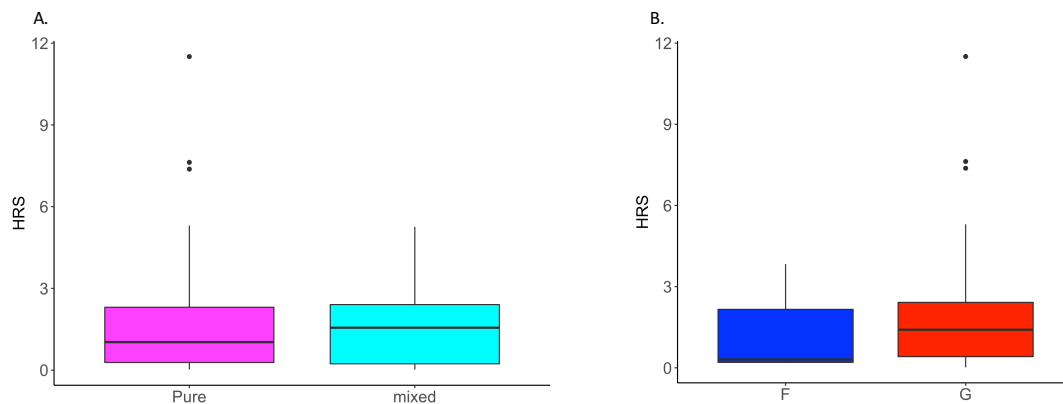


Figure C.11: A. The HRS scores of first generation offspring based status (pure vs mixed) B. The HRS scores of first generation offspring based on the mother’s population origin for n=67 mice.

Finally, a model was made for SI (Table C.3) initially including the same factors as in HRS and AA. The SI scores of offspring were influenced by sex (Figure C.12.A) whereby males had higher SI than females. Status (pure vs mixed) was also important, with pure individuals have slightly higher scores on average compared to the mixed mice (Figure C.12.B)- Finally, the population of the mother influenced the SI score, where there was a greater distribution in the scores of mice with a CB/G mother, and overall mice with a MC/F mother had higher scores (Figure C.12.C).

Table C.3. A Linear mixed effects model for SI scores of offspring to their parents. In the simplest model (AIC -183.72), there was an effect of status and the population of the mother on the SI of offspring. Mother and enclosure were dropped as a random effect (variance was 0). The random effect of father influenced to score.

Fixed Effect	Estimate	SE	t value	P-value (Chi ²)
Intercept	0.893	0.037	23.94	
Sex	0.028	0.010	2.74	0.006
Status	0.118	0.023	5.02	<0.001
Population Mother	-0.154	0.022	-6.73	<0.001
Random Effects			R	
Father ID	0.019	0.14	0.94	
Residual	0.001	0.03		

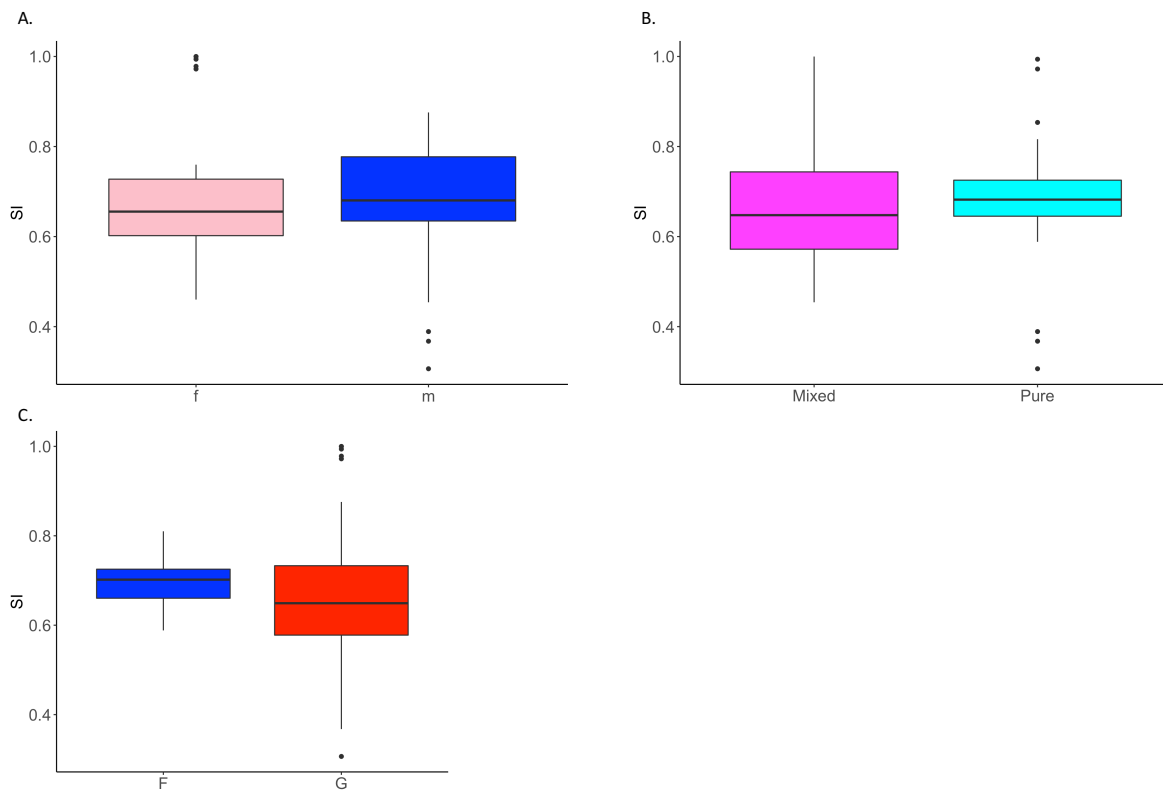


Figure C.12: A. The SI scores of first generation offspring based on sex B. The SI scores of first generation offspring based on status (pure vs mixed) C. The SI scores of first generation offspring based on the population origin of the mother for n=67 mice

Specific example of territory use

Looking at case studies of mouse families with stable territories, in our enclosures offspring stayed within or close to the territory of their fathers. In enclosure A the female mouse “Schuhmacher” produced offspring with males “Button” and “Kyvat”. Schumacher spent most of her time within the same territory as “Button”, “Kyvat’s” territory was adjacent. Offspring of “Schumacher” (male and female, Father “Button”: #250157, #250342, #250412, #250465, #250483, #250612, #251166, #251357 and #252800, Father “Kyvat”: #252305 and #254594) all used the territory of “Button” and “Kyvat” (Figure C.13).

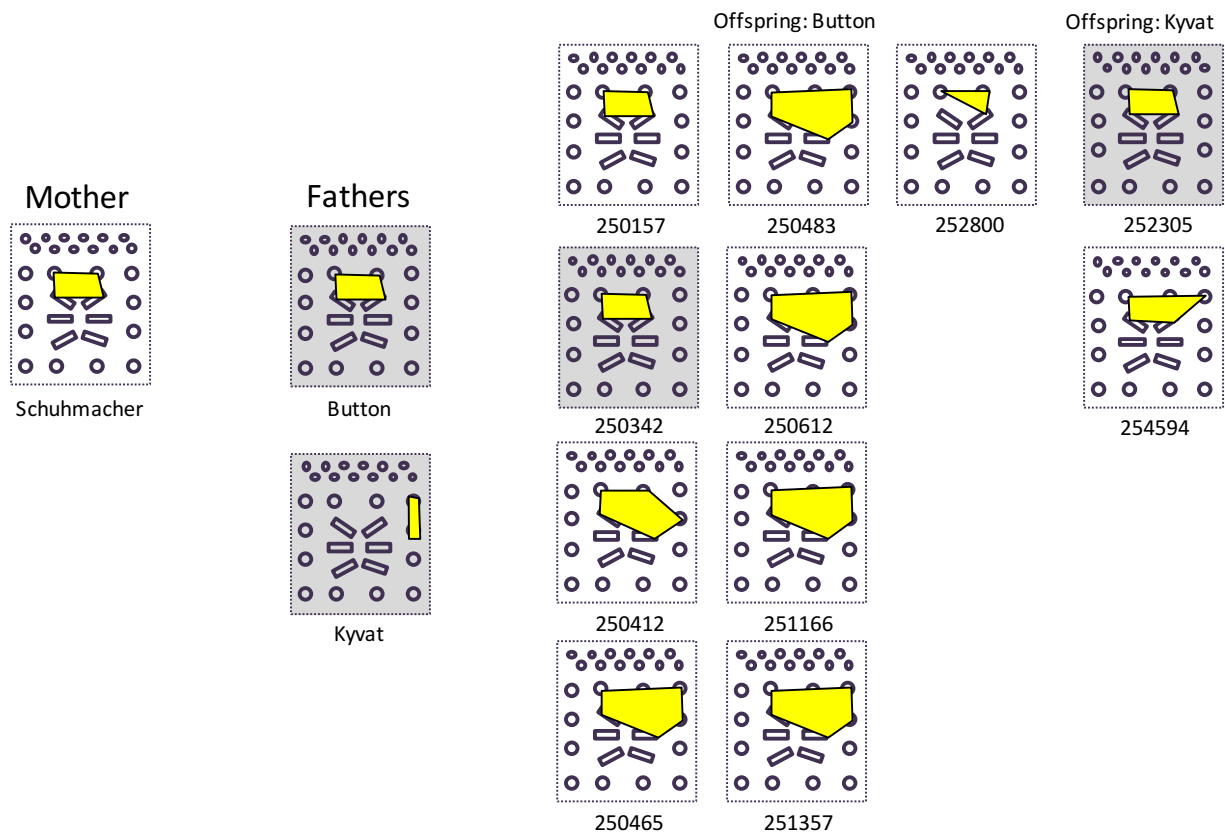


Figure C.13: Home ranges of the offspring of female mouse “Schuhmacher” and their respective fathers in enclosure A. Grey indicates that the offspring is male, white background indicates female mice. MCP indicates the area in which the mouse was found 95% of the time.

Male “Verstappen” produced offspring with females “Raikkonnen” and “Rosenburg”, “Raikkonnen” stayed within the same territory as “Verstappen”, while “Rosenburg” used the entire enclosure. Nevertheless, all of the offspring (“Raikkonnen”: #251164, #251218, #251250, #252945, #254483 and

#254841, “Rosenburg”: #253986 and #254810) remained within the territory of male “Verstappen” (See Figure C.14).

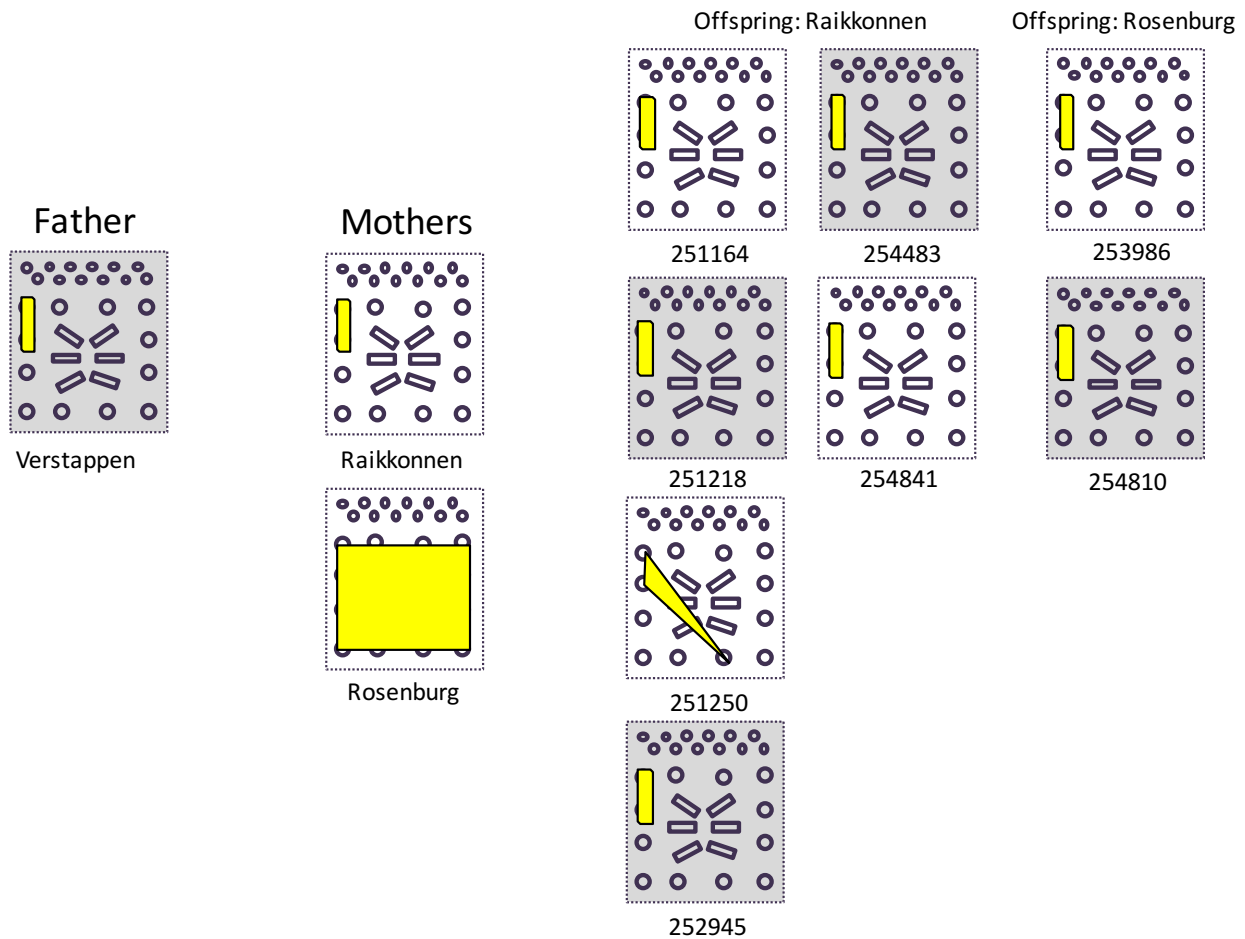


Figure C.14: Home ranges of the offspring of male mouse “Verstappen” and their respective mothers in enclosure A. Grey indicates that the offspring is male, the white background indicates female mice. MCP indicates the area in which the mouse was found 95% of the time.

In enclosure C the male “Subaru” (#250556) produced offspring with 3 females “Wiman” (#253668), “Piquet” (#254245) and “Isachsen” (#253408). “Subaru” spent most of his time in the front of the enclosure but produced offspring with “Piquet”, who was often at the back of the enclosure, on the same side. Offspring of “Piquet” and “Subaru” (#251798 and #254765) used both territories. The others spent most of their time in the territory of “Subaru” (#253362, #253819 and #254370). See Figure C.15.

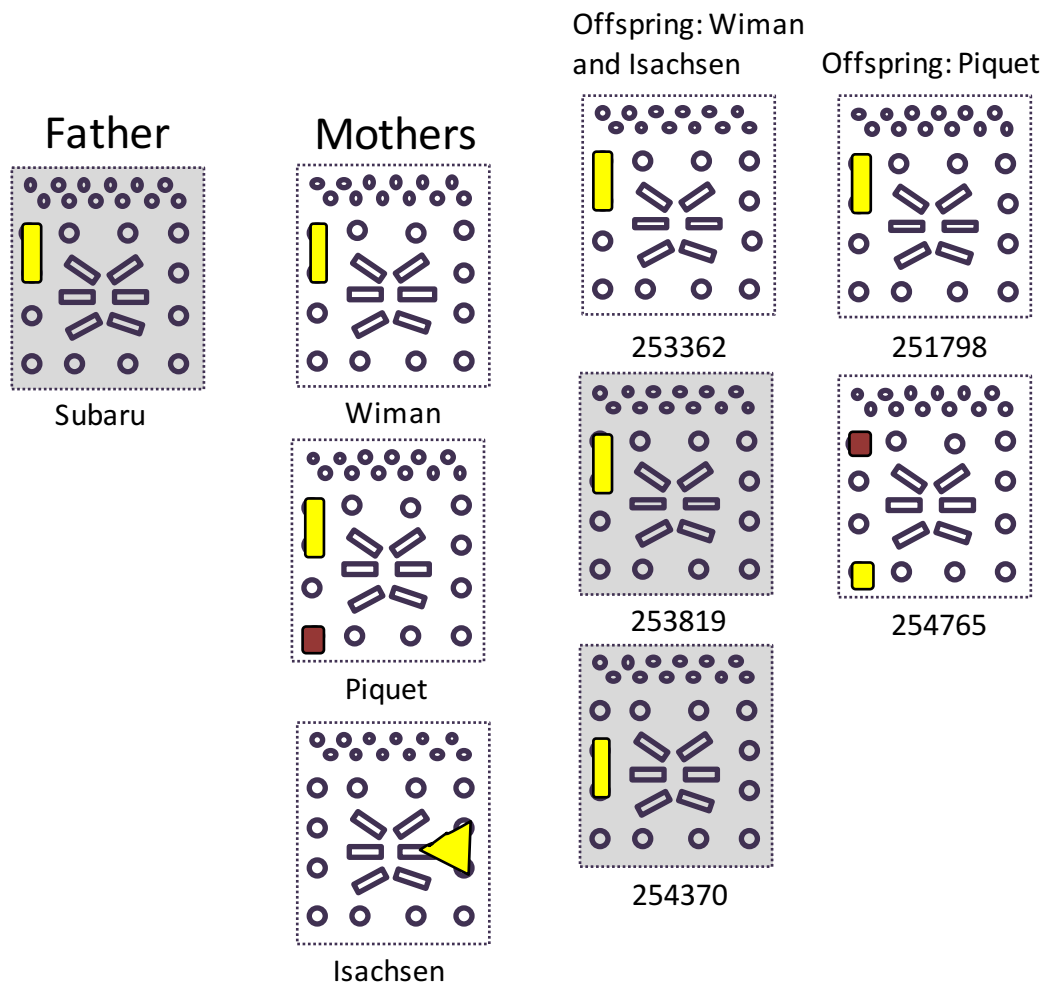


Figure C.15: Home ranges of the offspring of male mouse “Subaru” and their respective mothers in enclosure C. Grey indicates that the offspring is male, a white background indicates female mice. MCP indicates the area in which the mouse was found 95% of the time.

“Isachsen” lived primarily in the territory next to “Subaru”, which was occupied by male “Lasek”, who also produced 6 offspring with “Isachsen”. Offspring of “Isachsen” and “Lasek” never entered nest boxes occupied by the neighbour “Subaru” and spent most of their time in the territory of “Lasek” (#250491, #250675, #252533, #253020 and #254416). See Figure C.16.

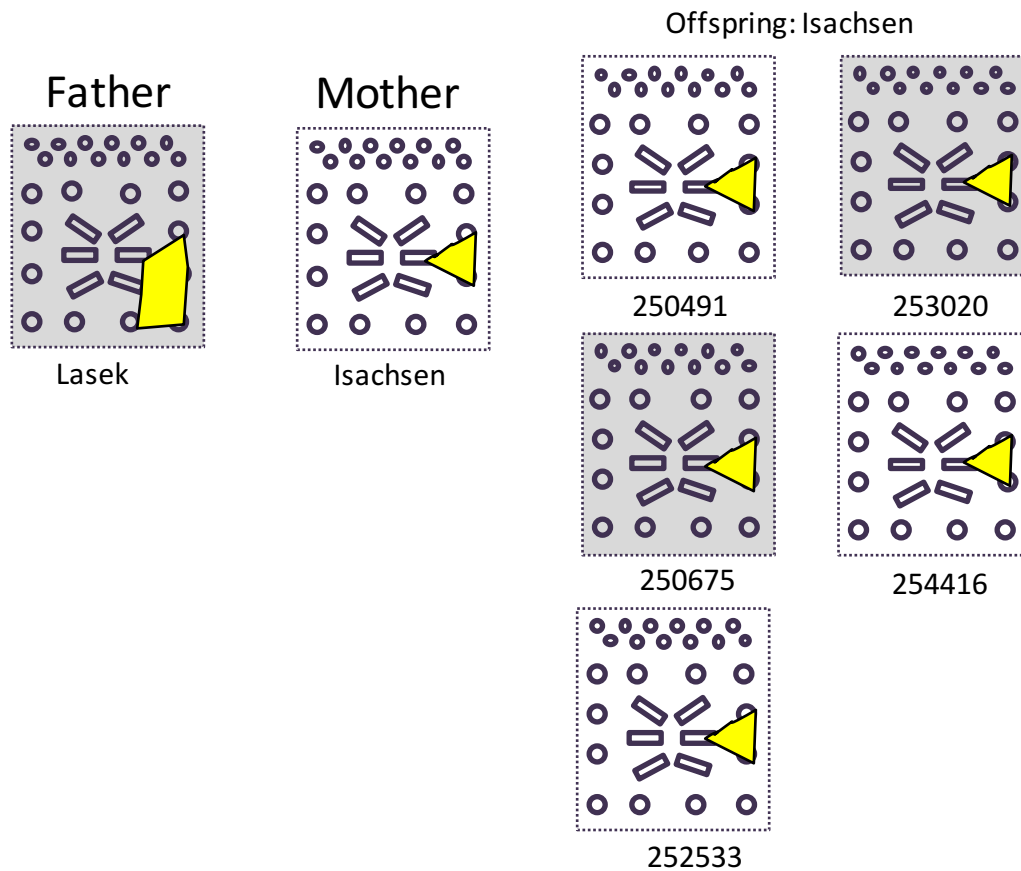


Figure C.16: Home ranges of the offspring of male mouse “Lasek” and “Isachsen” in enclosure C. The grey indicates that the offspring is male, the white background indicates female mice. MCP indicates the area in which the mouse was found 95% of the time.

D. Behavioural Tests

In order to measure personality types, a number of behavioural tests were carried out, including a Dark/Light box (DL), Startle Test (S) Novel Object test (NO), Open Field Test (OF) and an Elevated Plus Maze test (EP). For each of these test, multiple aspects were measured. These scores were checked for repeatability across time, clustered into related measurements and combined using a principle component analysis to make one final score for each individual.

Habituation

Dark/Light Box and Startle Test

The results from behavioural tests indicate that mice habituated to some of the experiments over time. In the case of the Dark/Light Box and Startle test (Figure D.1), mice from both populations were initially pooled to determine if there is an overall influence of repetition on the behavioural measurements. For this test, the number of mice tested decreased over time (Monitoring 1 n=52, 2 n=49 and 3 n=40). Kruskal-Wallis tests were carried out to determine the overall effect of time on the behavioural scores for each test. The first time the mouse poked its nose out of the shelter decreased over time (Figure D.1.A, Kruskal-Wallis test, chi-squared= 24.47, DF=2, p-value= <0.001). In contrast, the time it took for the mouse to emerge from the shelter was not significantly different (Figure D.1.B., Kruskal-Wallis test, chi-squared= 3.5089, DF=2, p-value= 0.173). After the startle, there were no significant differences between the time it took for the nose to emerge from the shelter (Figure D.1.C., Kruskal-Wallis test, chi-squared= 1.9118, DF=2, p-value= 0.385), and the time it took the mouse to leave the shelter (Figure D.1.D., Kruskal-Wallis test, chi-squared= 0.9227, DF=2, p-value= 0.6304).

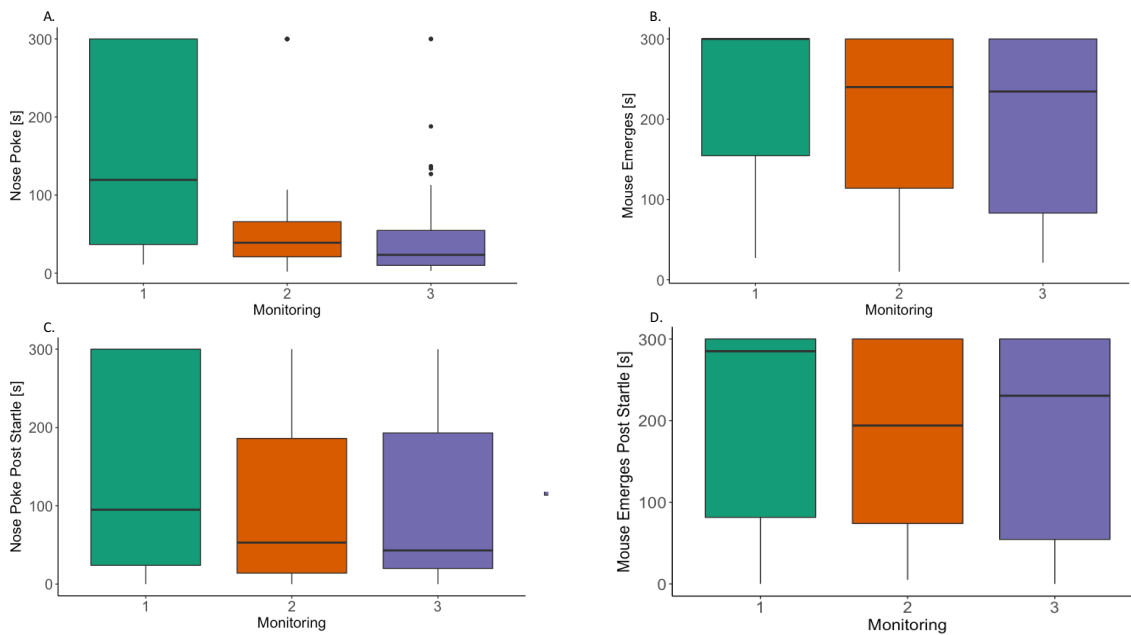


Figure D.1: Dark/Light Box test. A. The time it took for the mouse to look out of the shelter during the first five minutes. B. The time it took for the mouse to leave the shelter completely. C. After startle, the time it took for the mouse to look out of the shelter and D. the time it took for the mouse to emerge from the shelter post startle.

Novel Object Test

Mice also habituated to the novel object test (Figure D.2, Monitoring 1 n=46, 2 n=41 and 3 n=37). In this test, time spent with object (Figure D.2.A), visits to the object (Figure D.2.B), speed (Figure D.2.C) and sniff latency were measured (Figure D.2.D). Figure D.2.A. shows that overall the amount of time spent with the object was relatively consistent over time (Kruskal-Wallis test, chi-squared 2.954, DF=2, p-value=0.228) as was the number of visits made to the object (Figure D.2.B. Kruskal-Wallis test, chi-squared 0.971, DF=2, p-value=0.661). The measurements for speed (Figure D.2.C. Kruskal-Wallis test, chi-squared 6.49, DF=2, p-value=0.039) and sniff latency decreased over time (Figure D.2.D. Kruskal-Wallis test, chi-squared 18.86, DF=2, p-value=<0.001).

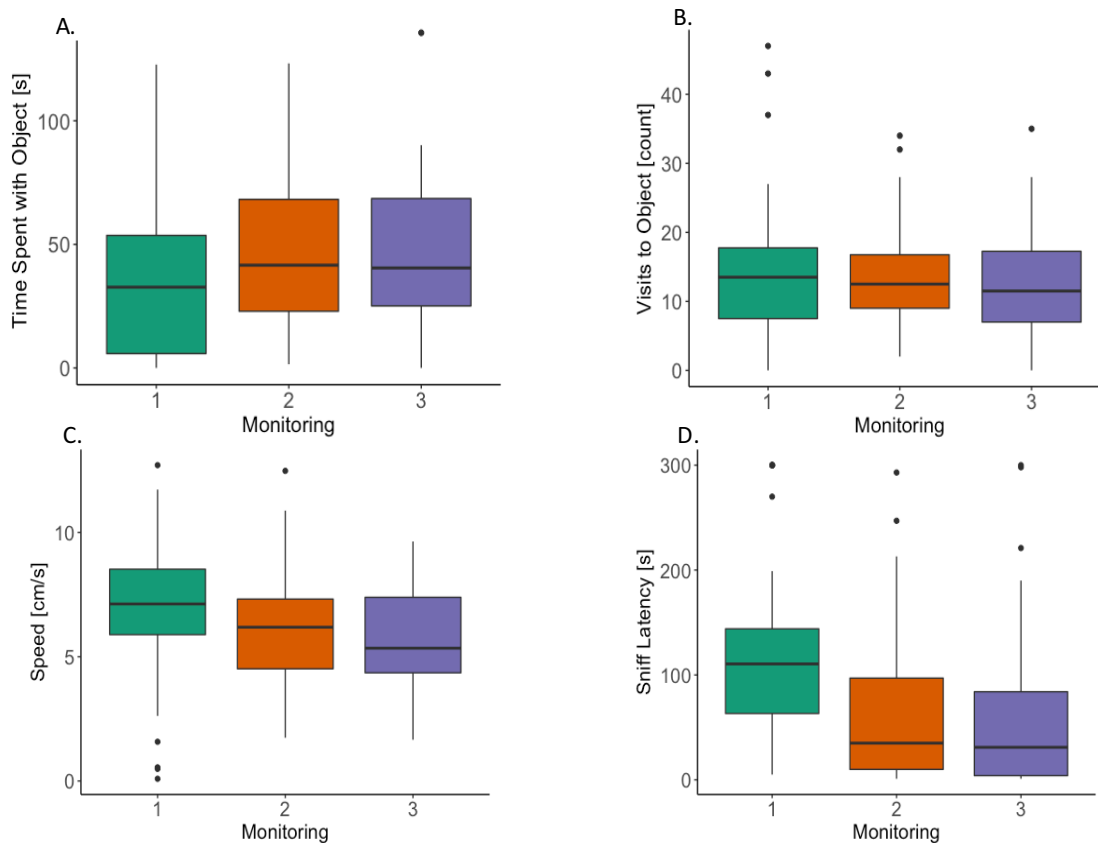


Figure D.2: Novel Object test. A. Time spent with the novel object. B. Visits to the novel object. C. The speed the mouse traveled during the novel object test and D. The sniff latency for mice by monitoring event.

Open Field Test

During the open field test, all three measurements changed over time (Monitoring 1 n=50, 2 n=48 and 3 n=40). Figure D.3.A. shows the amount of time the mouse spent at the wall over the course of three trials which changed over time (Kruskal-Wallis test, chi-squared= 10.769, DF=2, p-value=0.005). Figure D.3.B shows the distance traveled at the wall over all three measurements which changed significantly over time (Kruskal-Wallis test, chi-squared= 43.638, DF=2, p-value=<0.001). Finally, D.3.C. the speed of the mouse during the experiment also changed over time (Kruskal-Wallis test, chi-squared= 36.736, DF=2, p-value=<0.001).

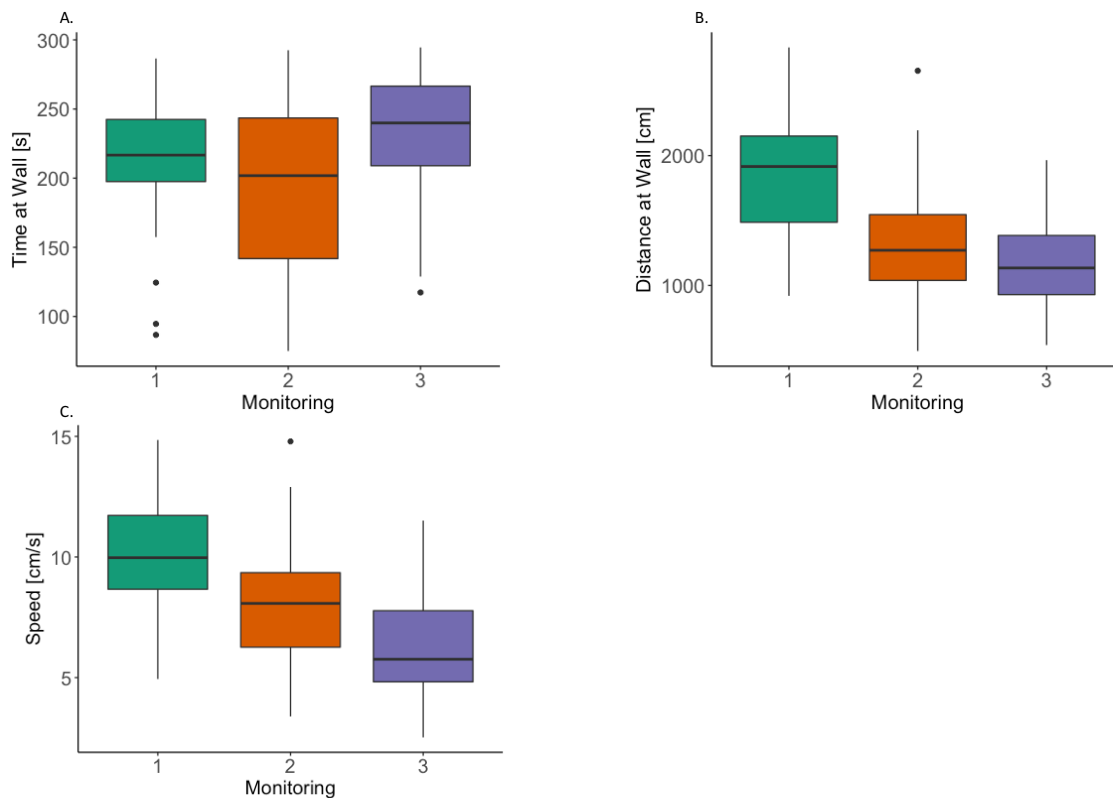


Figure D.3: Open Field Test. A. Time spent at the wall during the open field test. B. Distance travelled at the wall during open field test and C. The speed the mouse traveled during the open field test.

Elevated Plus Maze

Finally, the results for the elevated plus maze (Figure D.4, Monitoring 1 n=50, 2 n=48 and 3 n=40) The amount of time which the mouse spent in the light are during the elevated plus maze test did not change

over time (Figure D.4.A., Kruskal-Wallis test, chi-squared= 0.798, DF=2, p-value=0.671) while the speed the mouse traveled during the experiment decreased over time (Figure D.4.B., Kruskal-Wallis test, chi-squared= 23.606, DF=2, p-value=<0.001).

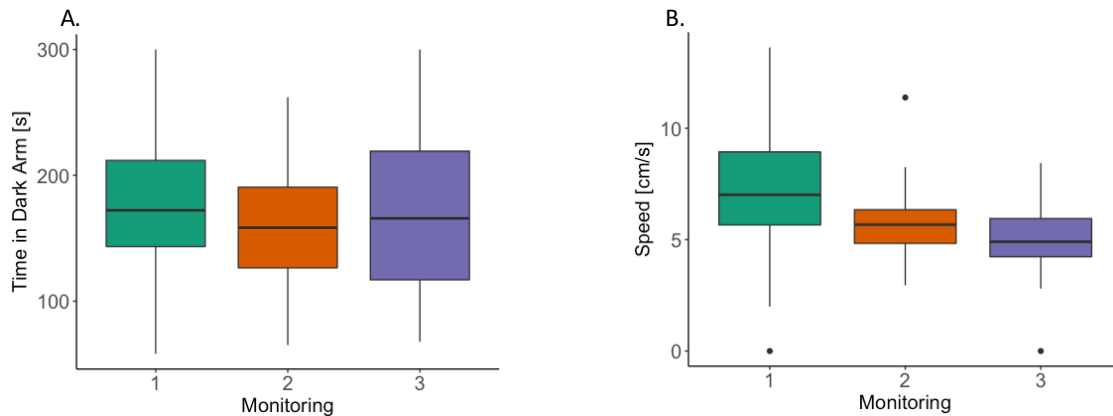


Figure D.4: Elevated plus maze. A. Time in Dark Arm during the elevated plus maze. B. The speed the mouse traveled during the elevated plus maze.

Controlling for missing individuals

To control for any differences between the three monitoring that may have resulted from the loss of individuals over time, statistical tests were repeated with just the mice who survived all three measurements (n=39 for Dark/Light box, Startle and Elevated Plus Maze and n=32 for Novel Object and Open Field). This did not change the results much in the dark/light box test nose poke (n=39, Kruskal-Wallis test, chi-squared= 17.176, DF=2, p-value=<0.001) or mouse leaving shelter (n=39, Kruskal-Wallis test, chi-squared= 1.6895, DF=2, p-value=0.4297); nor did it change the results of the startle test for the first nose poke post startle (n=39, Kruskal-Wallis test, chi-squared= 1.5919, DF=2, p-value=0.4511) and the first time the mouse left the shelter post startle (n=39, Kruskal-Wallis test, chi-squared= 0.561, DF=2, p-value=0.7552). During the novel object test, missing individuals did influence the overall changes in measurements of the time spent with the object (n=32, Kruskal-Wallis test, chi-squared= 0.167, DF=2, p-value=0.9197), the number of visits to the object (n=32, Kruskal-Wallis test, chi-squared= 2.2234, DF=2, p-value=0.329), speed (n=32, Kruskal-Wallis test, chi-squared= 9.3561, DF=2, p-value=0.009) and the

sniff latency (n=32, Kruskal-Wallis test, chi-squared= 13.43, DF=2, p-value=0.0012). This was the same for the open field test measurements for the time spent at the wall (n=35, Kruskal-Wallis test, chi-squared= 9.7811, DF=2, p-value=0.007), the distance travelled at the wall (n=35, Kruskal-Wallis test, chi-squared= 39.699, DF=2, p-value=0.001) and speed (n=35, Kruskal-Wallis test, chi-squared= 36.603, DF=2, p-value=0.001). Finally, in the elevated plus maze for time spent in the light arm (n=39, Kruskal-Wallis test, chi-squared= 1.0543, DF=2, p-value=0.5903) and the speed (n=39, Kruskal-Wallis test, chi-squared= 23.53, DF=2, p-value=<0.001).

Population origin influenced behaviour

Mice were then split by population and glmer/lmer models were fitted to determine whether there were overall differences in behavioural measurements between the two populations. In all models enclosure origin was kept as a random effect. Some data was transformed to fit models properly.

Dark/Light Box and Startle by population

Upon splitting the behavioural measurements by population, there was still a decrease overall in the length of time it took for the mouse to poke it's nose out of the shelter, but there was no interaction by population (Figure D.5.A, Table D.1). There was also an interaction between the time spent in the shelter and the population origin of the mouse in the time it took for the mouse to leave the shelter (Figure D.5.B, Table D.2). Mice also adapted to the Startle test over time, although (Figure D.5.C, Table D.3). The first time the mouse left the shelter was also influenced by population and monitoring (Figure D.5.D, Table D.4).

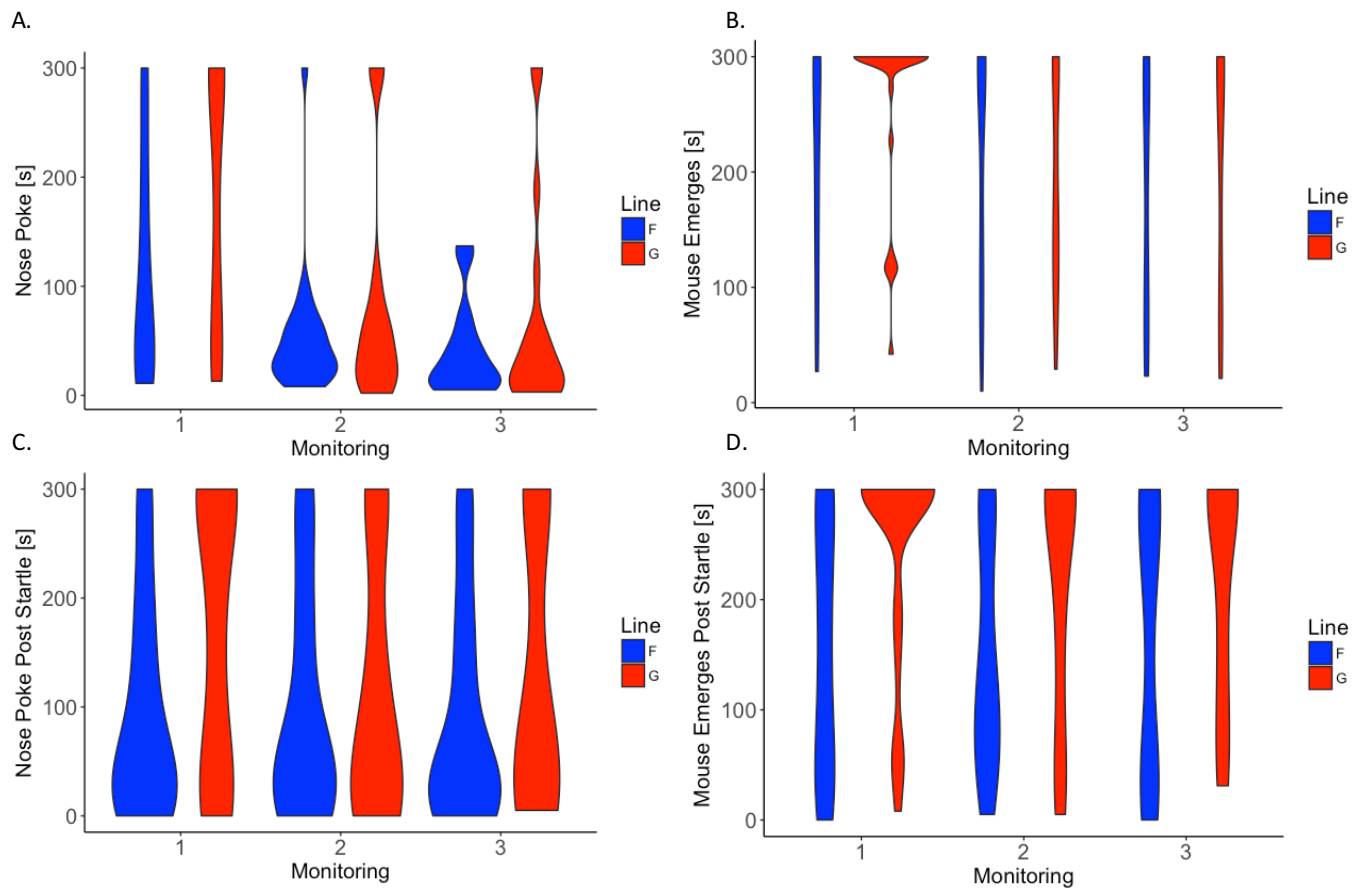


Figure D.5: Dark/Light Box test by population. A. The time it took for the mouse to look out of the shelter during the first five minutes. B. The time it took for the mouse to leave the shelter completely. C. After startle, the time it took for the mouse to look out of the shelter and D. the time it took for the mouse to emerge from the shelter post startle.

Table D.1: Fixed effects in the simplest model (AIC 218.42) including the random effects of enclosure and tester for nose poke outside shelter. Model fit using REML. Data transformed by log10 to ensure residuals fit a normal distribution.

	Estimate	SE	z value	P
Intercept	5.25	0.09	53.72	<0.001
Monitoring 2	-0.88	0.02	35.75	<0.001
Monitoring 3	-1.11	0.03	39.16	<0.001
Population	-0.49	0.02	27.81	<0.001

Table D.2: Fixed effects in the simplest model tested including enclosure and tester as random effects (AIC 8365.1). Model fit using maximum likelihood model. To fit this model a poisson distribution was used.

	Estimate	SE	z value	P-value
Intercept	5.52	0.10	50.31	<0.001
Monitoring 2	-0.33	0.02	-15.98	<0.001
Monitoring 3	-0.28	0.02	-12.69	<0.001
Population	-0.23	0.02	-12.42	<0.001
Population: Monitoring 2	0.30	0.03	11.16	<0.001
Population: Monitoring 3	0.10	0.03	3.64	<0.001

Table D.3: Fixed effects in the full model for nose poke after startle. Z-values and P-values for each effect were obtained from a series of likelihood ratio tests as described below. This model was compared to a simple model excluding the interaction term (Chi² 129.19, p-value <0.001) and fit the data significantly better therefore the full model was kept. Random effects were enclosure and tester.

	Estimate	SE	z value	P-value
Intercept	5.07	0.09	58.24	<0.001
Monitoring 2	-0.26	0.03	-9.70	<0.001
Monitoring 3	-0.10	0.03	-3.78	<0.001
Population	-0.59	0.03	-23.55	<0.001
Population: Monitoring 2	0.42	0.04	11.35	<0.001
Population: Monitoring 3	0.18	0.04	4.62	<0.001

Table D.4: Fixed effects in the full model for mouse to leave box after startle. Z-values and P-values for each effect were obtained from a series of likelihood ratio tests as described below. This model was compared to a simple model excluding the interaction term (Chi² 39.13, p-value <0.001) and fit the data significantly better therefore the full model was kept. Random effects were enclosure and tester.

	Estimate	SE	z value	P-value
Intercept	5.48	0.07	82.04	<0.001
Monitoring 2	-0.06	0.02	-3.02	0.002
Monitoring 3	-0.05	0.02	-2.28	0.022
Population	-0.44	0.02	-22.45	<0.001
Population: Monitoring 2	0.10	0.03	3.60	<0.001
Population: Monitoring 3	0.18	0.03	6.18	<0.001

Mice also habituated to the novel object test (Figure D.6). In this test, time spent with object, visits to the object, speed and sniff latency were measured. Speed was normally distributed; the other three measures were transformed using square root +1 and a linear mixed effect model with random effects of enclosure and tester were used on all four measurements. The time spent with object was influenced by population and monitoring (Figure D.6.A, Table D.5). This was also the case for the visits to the object (Figure D.6.B, Table D.5), and speed (Figure D.6.C, Table D.5). In the case of sniff latency, there was an overall influence of monitoring and population (Figure D.6.D, Table D.5).

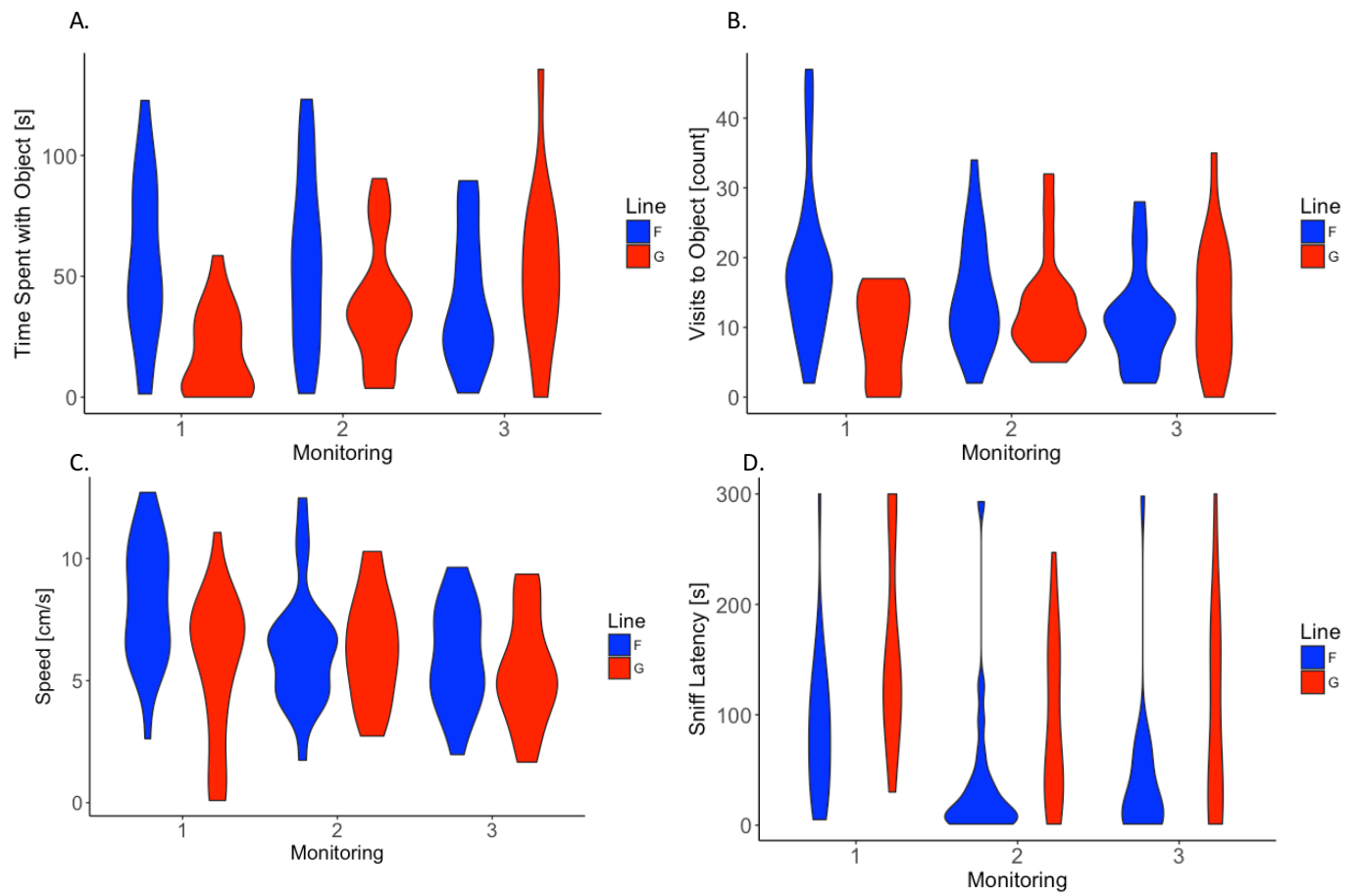


Figure D.6: Novel Object test by population. A. Time spent with the novel object. B. Visits to the novel object. C. The speed the mouse traveled during the novel object test and D. The sniff latency for each monitoring event.

Table D.5: Fixed effects for the simplest model (lowest AIC) for novel object measurements. T-values and P-values for each effect were obtained from a series of likelihood ratio tests as described below. Random effects were enclosure and tester. Time spent with object and visits to object were transformed with square root + 1, to correct left skew. Sniff latency models were made using glmer (Poisson), model was selected based on lowest AIC score.

Time Spent with Object	Estimate	SE	T value (z)	P value
Intercept	3.84	0.52	7.28	<0.001
Monitoring 2	2.05	0.75	2.72	0.007
Monitoring 3	2.96	0.78	3.76	<0.001
Population	3.27	0.70	4.61	<0.001
Monitoring 2: Population	-2.01	1.03	-1.95	0.0533
Monitoring 3: Population	-3.99	1.07	-3.73	<0.001
Visits to the Novel Object				
Intercept	2.97	0.23	12.68	<0.001
Monitoring 2	0.69	0.34	2.01	0.046
Monitoring 3	0.63	0.36	1.75	0.084
Population	1.30	0.32	4.01	0.001
Monitoring 2: Population	-1.31	0.47	-2.41	0.017
Monitoring 3: Population	-1.51	0.49	-3.09	0.003
Speed				
Intercept	5.69	0.52	10.96	<0.001
Monitoring 2	0.65	0.76	0.85	0.395
Monitoring 3	-0.28	0.80	-0.34	0.730
Population	2.55	0.73	3.51	<0.001
Monitoring 2: Population	-2.59	1.04	-2.48	0.014
Monitoring 3: Population	-2.00	1.09	-1.82	0.070
Sniff Latency				
Intercept	5.01	0.07	75.69	<0.001
Monitoring 2	-0.44	0.04	-11.19	<0.001
Monitoring 3	-0.38	0.03	-11.84	<0.001
Population	-0.60	0.03	-22.39	<0.001
Monitoring 2: Population	-0.29	0.05	-5.87	<0.001
Monitoring 3: Population	-0.31	0.05	-6.02	<0.001

The measurements taken in the open field test (Monitoring 1 n=46, 2 n=41 and 3 n=37) were also subject to habituation (Figure D.7). A linear mixed effects model was used, including the random effects of tester present in enclosure and enclosure itself. Time travelled at the wall (Figure D.7.A, Table D.6), distance travelled at the wall (Figure D.7.B, Table D.6) and speed (Figure D.7.C, Table D.6) were also influenced by time and population origin, overall measurements decreased over time, but measurements for MC/F mice were higher than CB/G mice.

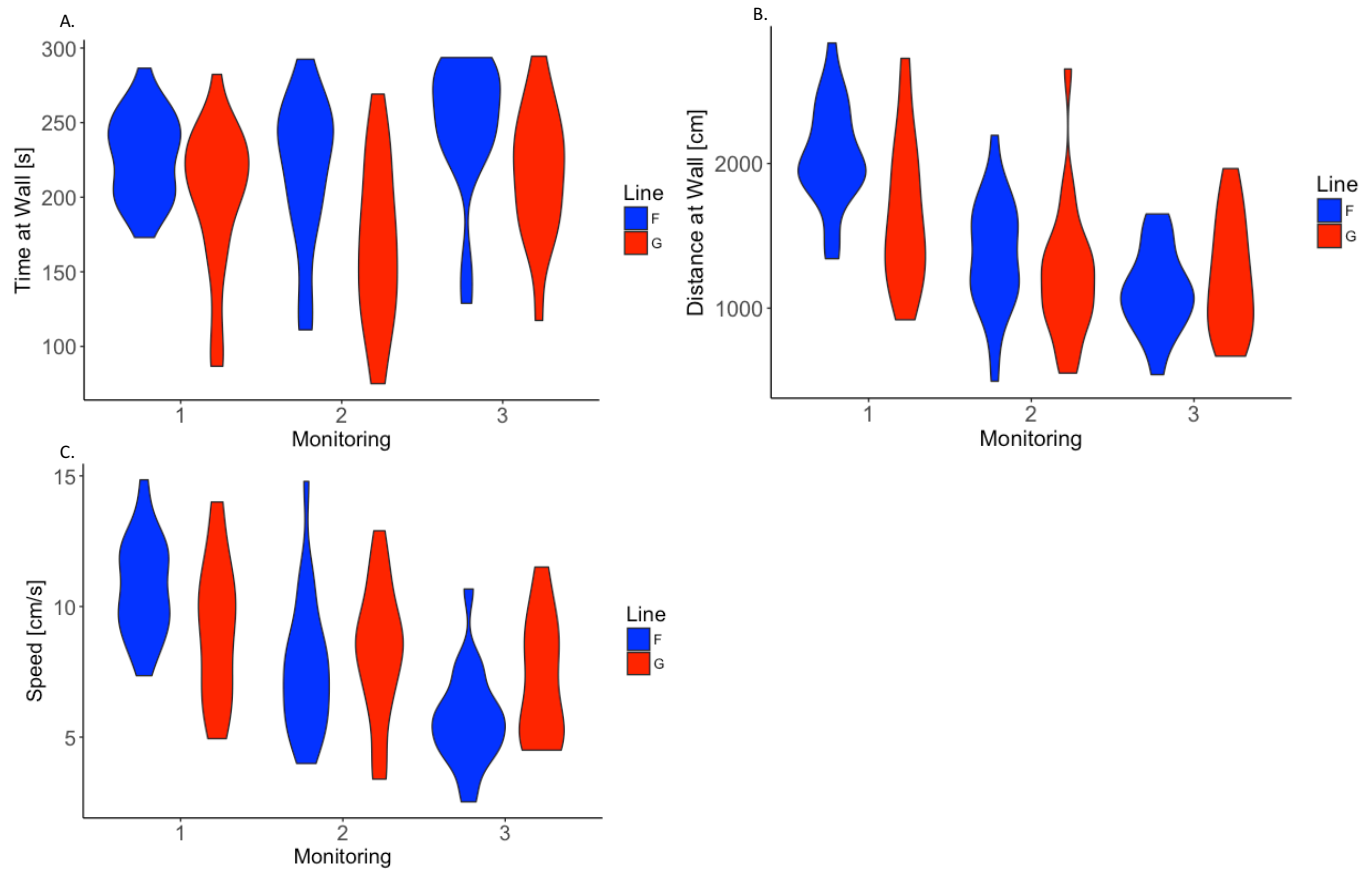


Figure D.7: Open Field Test by population. A. Time spent at the wall during the open field test. B. Distance travelled at the wall during open field test and C. The speed the mouse traveled during the open field test.

Table D.6: Fixed effects in the simplest model (lowest AIC) for the behavioural traits measured in open field test. T-values and P-values for each effect were obtained from a series of likelihood ratio tests as described below. Random effects were enclosure and tester.

Time at Wall	Estimate	SE	T value	P-value
Intercept	202.98	10.15	20.00	<0.001
Monitoring 2	-36.56	14.04	-2.60	0.010
Monitoring 3	13.67	14.76	0.92	0.356
Population	26.41	13.01	2.02	0.045
Monitoring 2: Population	24.62	18.81	1.30	0.194
Monitoring 3: Population	6.34	19.86	0.31	0.750
Distance (Wall)				
Intercept	1670.30	84.53	19.52	<0.001
Monitoring 2	-417.33	122.36	-3.41	0.009
Monitoring 3	-437.06	129.67	-3.37	0.001
Population	403.94	117.13	3.44	0.001
Monitoring 2: Population	-269.12	169.39	-1.59	0.115
Monitoring 3: Population	-498.94	178.99	-2.79	0.006
Speed				
Intercept	9.15	0.48	19.40	<0.0001
Monitoring 2	-0.91	0.69	-1.31	0.1915
Monitoring 3	-1.96	0.73	-2.66	0.0088
Population	1.56	0.66	2.35	0.0202
Monitoring 2: Population	-2.15	0.96	-2.24	0.0269
Monitoring 3: Population	-3.05	1.02	-2.99	0.0034

Finally, the results for the elevated plus maze (Figure D.8, M1 n=50, M3 n=48 and M5 n=40) were also calculated using a linear mixed effects model with random effects of Tester and Enclosure. The time spent in the dark arm was normally distributed; speed was transformed using square root ($x + 1$) transformation (Figure D.8.A). There was an overall effect of population but not monitoring (no habituation in this case) (Table D.7). In the case of speed in the elevated plus maze, there was no difference between the populations, only monitoring (Figure D.8.B).

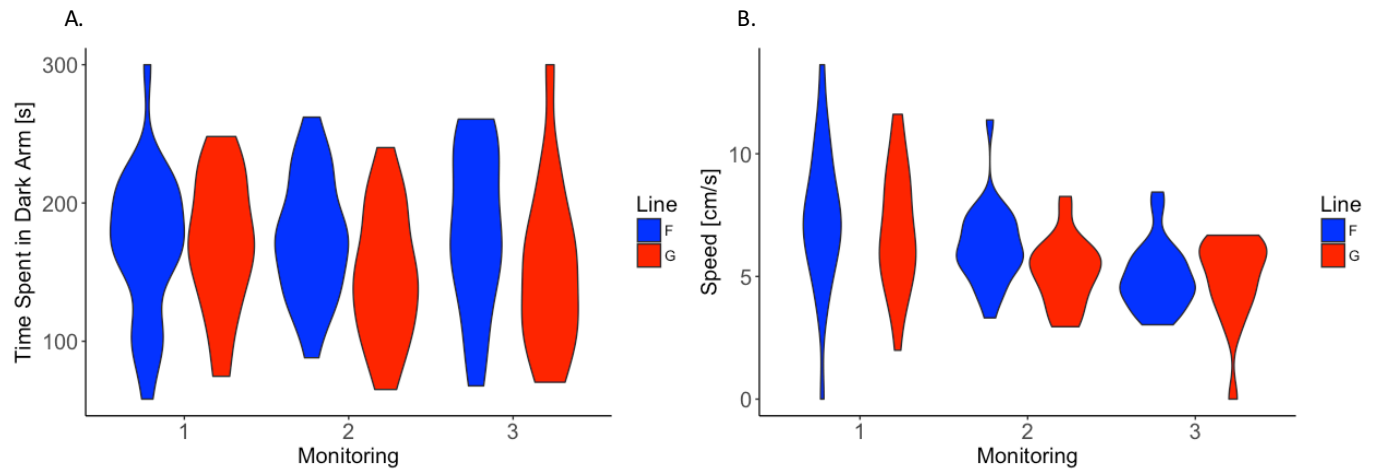


Figure D.8: Elevated plus maze by population. A. Time in Dark Arm during the elevated plus maze. B. The speed the mouse traveled during the elevated plus maze.

Table D.7: Fixed effects in the simplest model (lowest AIC) for behavioural measurements in the elevated plus maze. T-values and P-values for each effect were obtained from a series of likelihood ratio tests as described below.

Time in Dark Arm	Value	S.E.	T value	P
Intercept	170.53	10.43	16.34	<0.001
Monitoring 2	-22.57	14.91	-1.51	0.1328
Monitoring 3	-19.10	16.20	-1.17	0.2406
Population	-0.67	14.47	-0.05	0.9627
Monitoring 2: Population	27.05	20.68	1.31	0.1931
Monitoring 3: Population	29.71	21.83	1.36	0.1760
Speed				
Intercept	1.98	0.06	34.34	<0.001
Monitoring 2	-0.15	0.07	-2.12	0.035
Monitoring 3	-0.29	0.07	-4.00	<0.001
Population	0.08	0.06	1.41	0.1599

Controls Mice Habituation

Dark/Light Box

There was a significant difference between the first and last measurement of the first time the mouse looked out of the shelter for control group one (Figure D.9.A, Paired Wilcoxon Test, N=11, V=52, p-value = 0.014) but not for control group two (Figure D.9.A, Paired Wilcoxon Test, N=11, V=28, p-value = 1) nor was there a difference between the first and last measurement from control group one (Figure D.9.B,

Paired Wilcoxon Test, $V=36$, $N=11$, $p\text{-value} = 0.123$) and control group two (Figure D.9.B, Paired Wilcoxon Test, $V=25$, $N=11$, $p\text{-value} = 0.845$) for the first time the mouse emerged from the shelter completely.

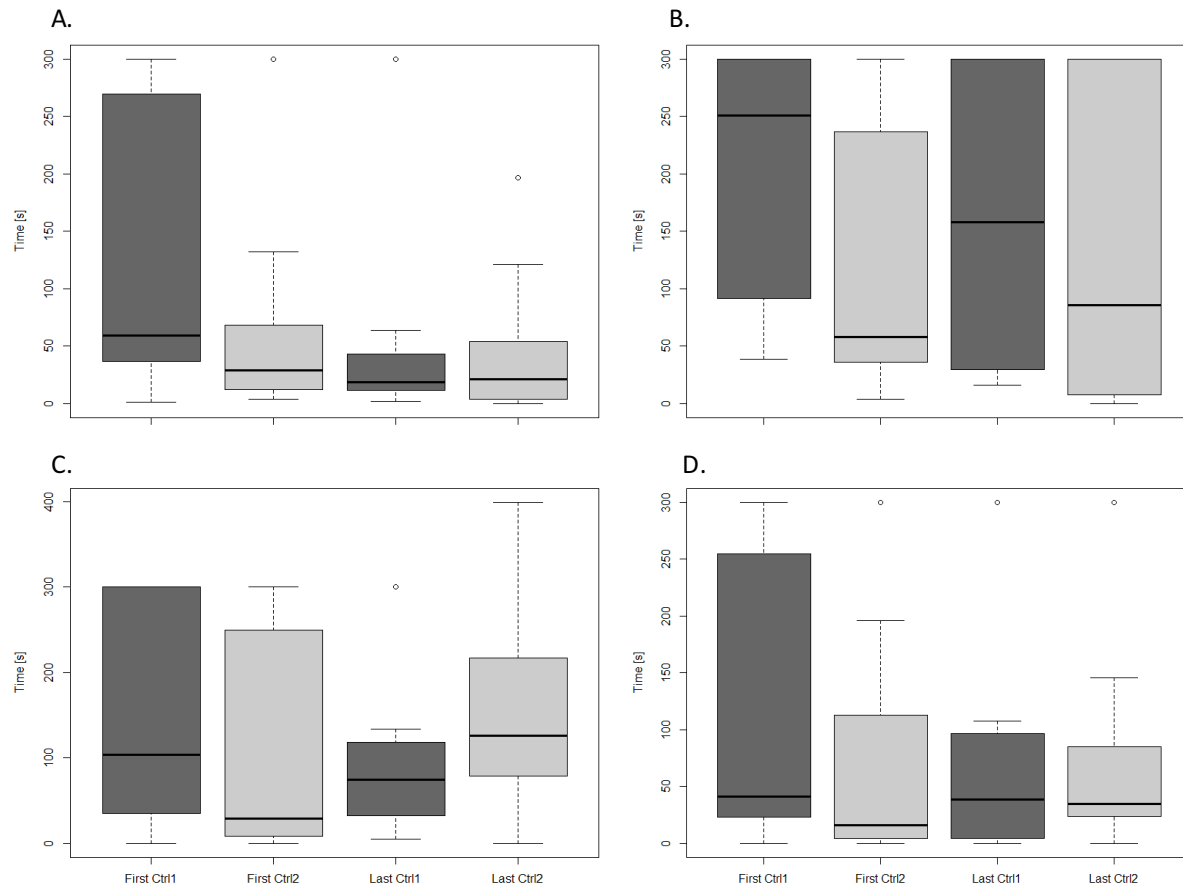


Figure D.9: Dark/Light box for control mice. A. Time in seconds it took for the mouse to look out of the shelter during the first and last measurement from control group one and control group two. B. Time for the mouse to emerge completely from the shelter in control group one and two, for the first and last measurement. C. time in seconds it took for the mouse to look out of the shelter post startle during the first and last measurement from control group one and control group two. D. Time in seconds it took for the mouse to emerge completely from the shelter post startle in control group one and two, for the first and last measurement.

Startle Test

Post startle, there was no significant difference between the first and last measurement of the first time the mouse looked out of the shelter for control group one (Figure D.9.C, Paired Wilcoxon Test, $V=31$, $N=11$, $p\text{-value} = 0.759$) or two (Figure D.9.C, Paired Wilcoxon Test, $V=17$, $N=11$, $p\text{-value} = 0.554$) nor was

there a difference between the first and last measurement from control group one (Figure D.9.D, Paired Wilcoxon Test, $V=36$, $N=11$, $p\text{-value} = 0.415$) and control group two (Figure D.9.D, Paired Wilcoxon Test, $V=17$, $N=11$, $p\text{-value} = 0.553$) for the first time the mouse emerged from the shelter completely.

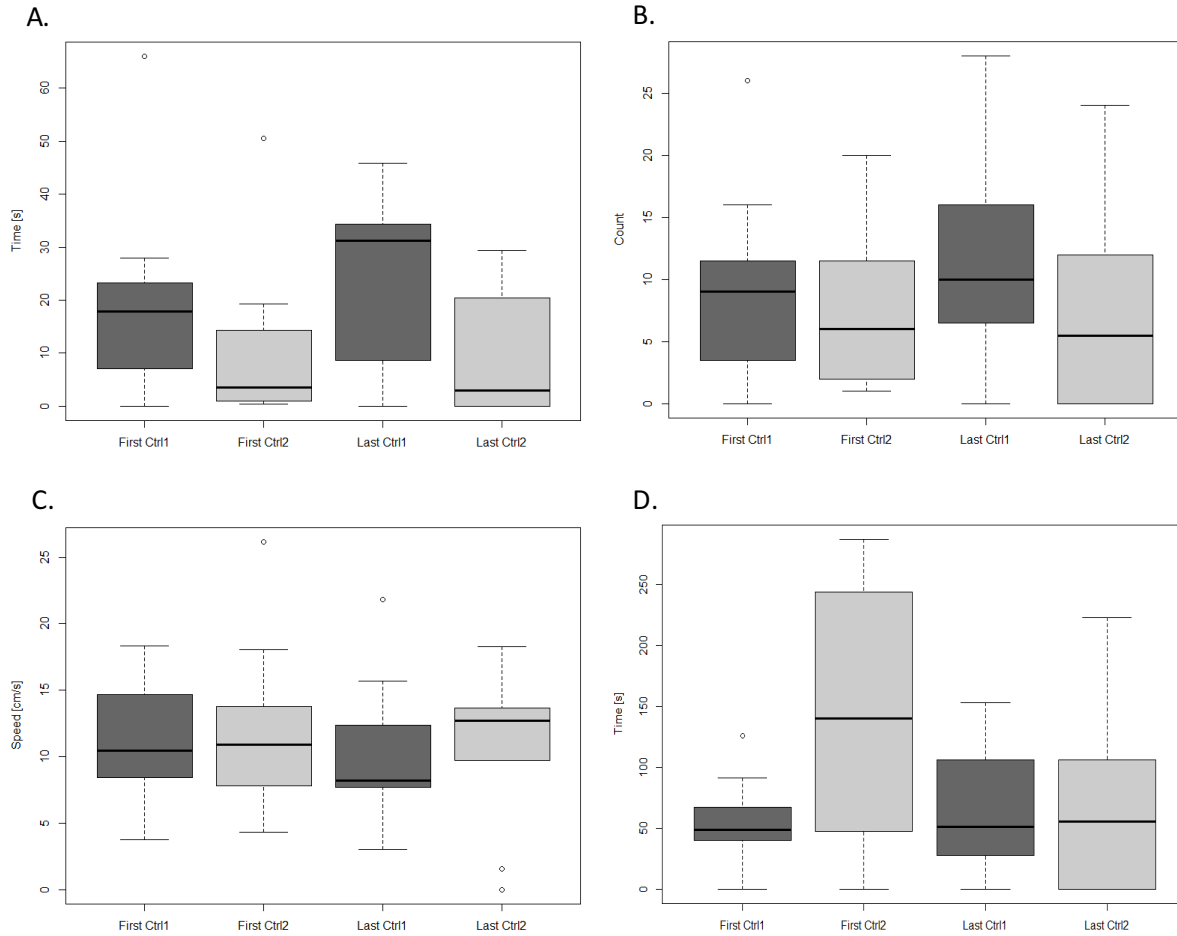


Figure D.10: Novel Object Test for control mice. A. Time with object in seconds by control group and first and final measurements B. Count of visits to object by control group and first and final measurements C. Speed in cm/s by control group and first and final measurements and D. Sniff latency by control group and first and final measurements.

There was no significant difference between the first and last measurement of time spent with the novel object for control group one (D.10.A, Paired Wilcoxon Test, $V=17$, $N=11$, $p\text{-value} = 0.554$) or two (D.10.A Paired Wilcoxon Test, $V=22$, $N=11$, $p\text{-value} = 0.625$). This was the same for the number of visits in control group one (D.10.B, Paired Wilcoxon Test, $V=12.5$, $N=11$, $p\text{-value} = 0.483$) and two (D.10.B, Paired Wilcoxon Test, $V=19$, $N=11$, $p\text{-value} = 0.7199$), speed in the novel object test in control group one (D.10.C, Paired Wilcoxon Test, $V=36$, $N=11$, $p\text{-value} = 0.432$) and two (D.10.C, Paired Wilcoxon Test,

V=37, N=11, p-value = 0.375) and the sniff latency for control group 1 (D.10.D, Paired Wilcoxon Test, V=19, N=11, p-value = 0.722) and 2 (D.10.D, Paired Wilcoxon Test, V=37, N=11, p-value = 0.097).

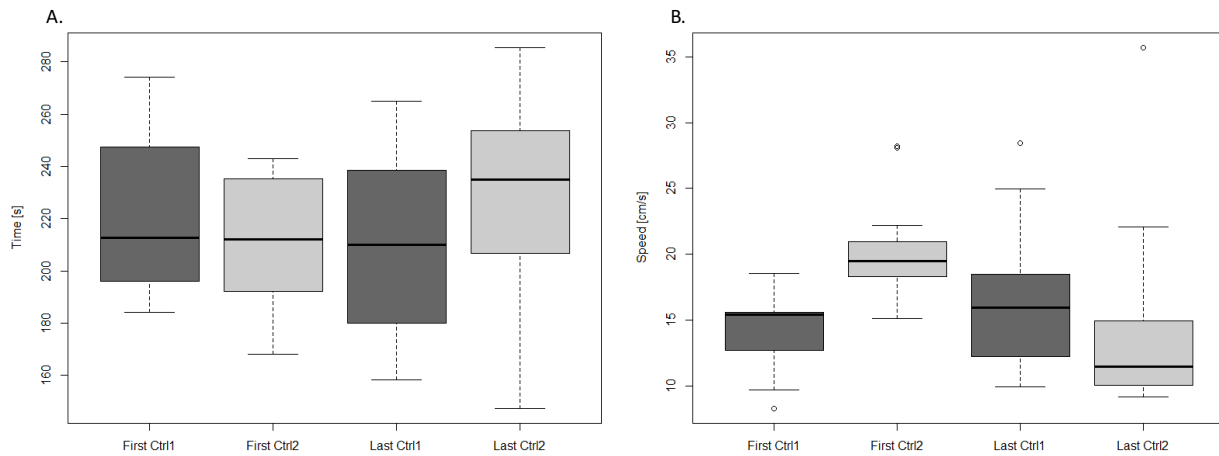


Figure D.11: Open Field Test for control mice. A. Time in seconds it took for the mouse spent at the wall during the first and last measurement from control group one and control group two. B. speed of the mouse in the open field test by control group one and two, for the first and last measurement.

There was no significant difference between the first and last measurement of time spent at the wall for control group one (Figure D.11.A, Paired Wilcoxon Test, V=35, N=11, p-value = 0.898) or two (Figure D.11.A, Paired Wilcoxon Test, V=15, N=11, p-value = 0.232). The speed in the open field did not differ between the first and last measurement from control group one (Figure D.11.B, Paired Wilcoxon Test, V=18, N=11, p-value = 0.206), while the first and last measurements from control group two were significantly different over time (Figure D.11.B, Paired Wilcoxon Test, V=48, N=11, p-value = 0.037).

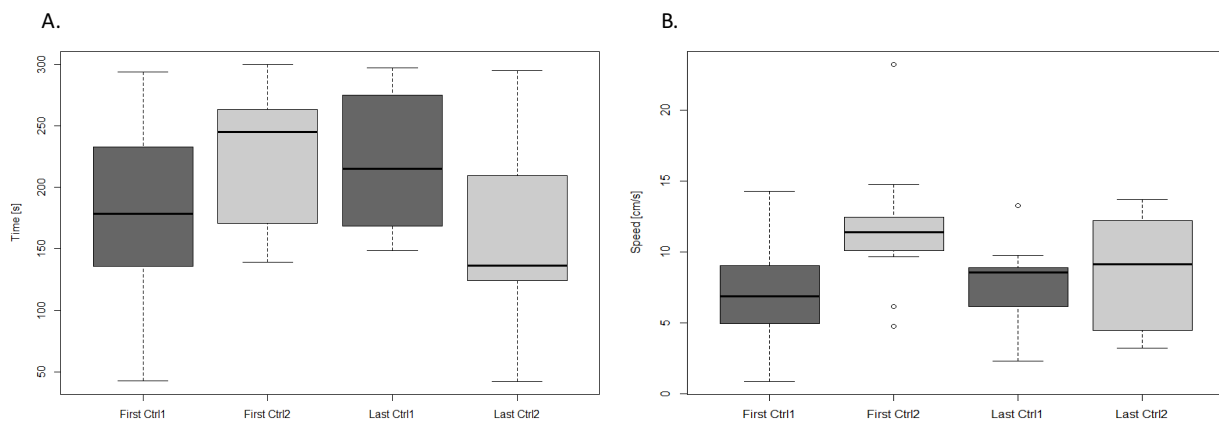


Figure D.12: Elevated Plus Maze for control mice. A. Time in seconds it took for the mouse spent in the dark arm during the first and last measurement from control group one and control group two. B. speed

of the mouse in the elevated plus maze test by control group one and two, for the first and last measurement.

There was no significant difference between the first and last measurement of control group one (Figure D.12.A, Paired Wilcoxon Test, $V=17$, $N=11$, $p\text{-value} = 0.168$) or two (Figure D.12.A, Paired Wilcoxon Test, $V=43$, $N=11$, $p\text{-value} = 0.13$) in the amount of time spent in the dark arm. This was the same for speed in control group one (Figure D.12.B, Paired Wilcoxon Test, $V=19$, $N=11$, $p\text{-value} = 0.722$) and two (Figure D.12.B, Paired Wilcoxon Test, $V=43$, $N=11$, $p\text{-value} = 0.13$).

Repeatability

As the repeatability of a behaviour is a key component for the identification of personality traits, all single measurements assessed in the personality tests were tested investigated using rptR and ICC (Table D.8). Some measurements were moderately repeatable during the course of the experiment ($R > \text{or} = 0.30$), this included measurements for the mouse leaving the shelter during the dark/light box test, and speed and sniff latency in the novel object test. Others were weakly repeatable ($R < 0.30$) such as nose poke and mouse appearance after startle, time spent with object in the novel object test, time spent at the wall in the open field test and time spent in the dark arm in the elevated plus maze. Other measurements were not repeatable (low R or non-significant $p\text{-value}$), these included distance along the wall and speed in the open field test and speed in the elevated plus maze. To more generally describe whole personality traits and not only single measurements, we performed a Cluster correlation analysis (see next section "Cluster correlation").

Table D.8: Repeatability calculated using rptR (linear mixed models, $\text{Repeatability} = \text{Var}_{\text{Individual}} / (\text{Var}_{\text{Individual}} + \text{Var}_{\text{Residual}})$). P-values are calculated using LRT. ICC scores made using MCMC estimate in ICC package (R).

Measure	Definition	Mean SE	N	R	P	ICC
Nose Poke	First time mouse looks out of box (s)	0.08	141	0.09	0.169	0.25
Mouse Appears	First time mouse leaves box (s)	0.09	141	0.35	0.0001	0.27
Nose Poke	First time mouse looks out after startle (s)	0.10	141	0.19	0.036	0.22
Mouse Appears	First time mouse comes out after startle (s)	0.07	141	0.19	0.017	0.21
Time with Object	Time spent visiting novel object (s)	0.10	124	0.24	0.001	0.24
Visits to Object	Number of visits to novel object (count)	0.11	124	0.27	0.005	0.25
Speed (Novel Object)	Speed (cm/s)	0.10	124	0.31	0.003	0.30
Sniff Latency	First Visit to object (s)	0.11	126	0.31	0.002	0.41
Wall Time	Time spent at wall (s)	0.10	124	0.23	0.0008	0.24
Wall Distance	Distance travelled at wall (cm)	0.07	124	0.04	0.009	0.06
Speed (Open Field Test)	Speed (cm/s)	0.09	124	0.10	1	0.12
Dark Arm	Time spent in Arm Dark (s)	0.10	138	0.18	0.0023	0.22
Speed (Elevated Plus Maze)	Speed (cm/s)	0.19	138	0.23	0.0744	0.21

Cluster correlation

Besides repeatability, the consistency of behavioural reactions in different situations is needed to define an animals' personality. Only individuals which were measured three times were included in the analysis (n=33). We estimated this consistency using a spearman correlation matrix (Figure D.13, full correlation matrix with Holm corrected p-values attached in supplemental information, Figure S4). We found positive correlations between the time it took the mouse to leave the shelter, the first time the nose poked out of the shelter and the speed travelled in the novel object test and negative correlations between these and the time spent in with the novel object. Further, we performed a cluster analysis to group measurements provoking the same reactions of individuals in different contexts (Figure D.14). The first cluster shows variables commonly associated with activity related measurements (such as sniff latency (NO), Speed (NO), the number of visits to the Novel Object (NO) and the amount of time spent with the Novel Object (NO)). The second, anxiety, using measurements from the open field test, elevated plus maze test and dark/light box commonly associated with anxiety (mouse emerges from shelter (DL), time spent in dark arm (EP) and time spent at wall (OF)). Finally, the third category was named curiosity/risk taking, both measurements from the startle test, the first nose poke (S) and the first time the mouse left the shelter (S). These categories were used to create behavioural scores which will be used to access fitness of behavioural types below.

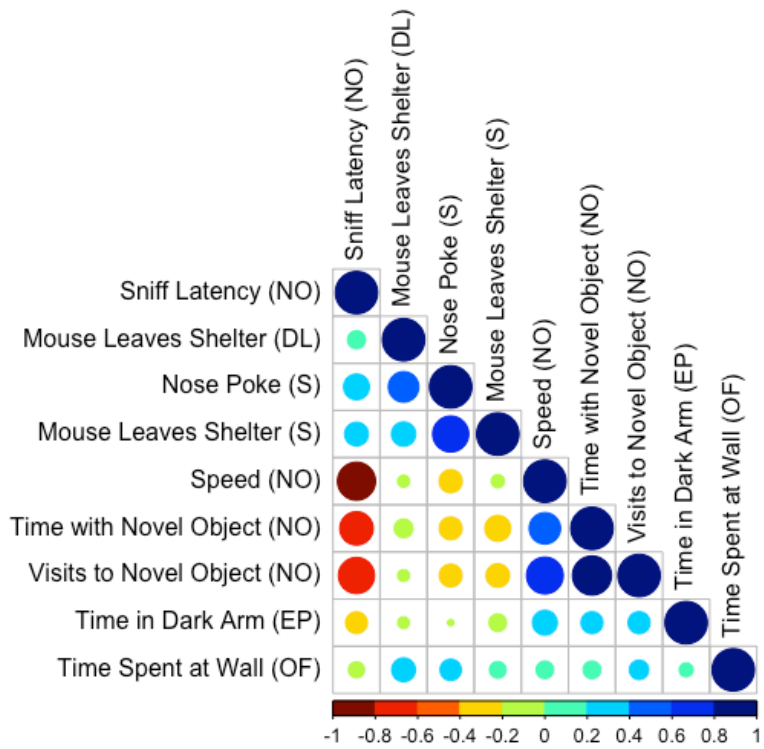
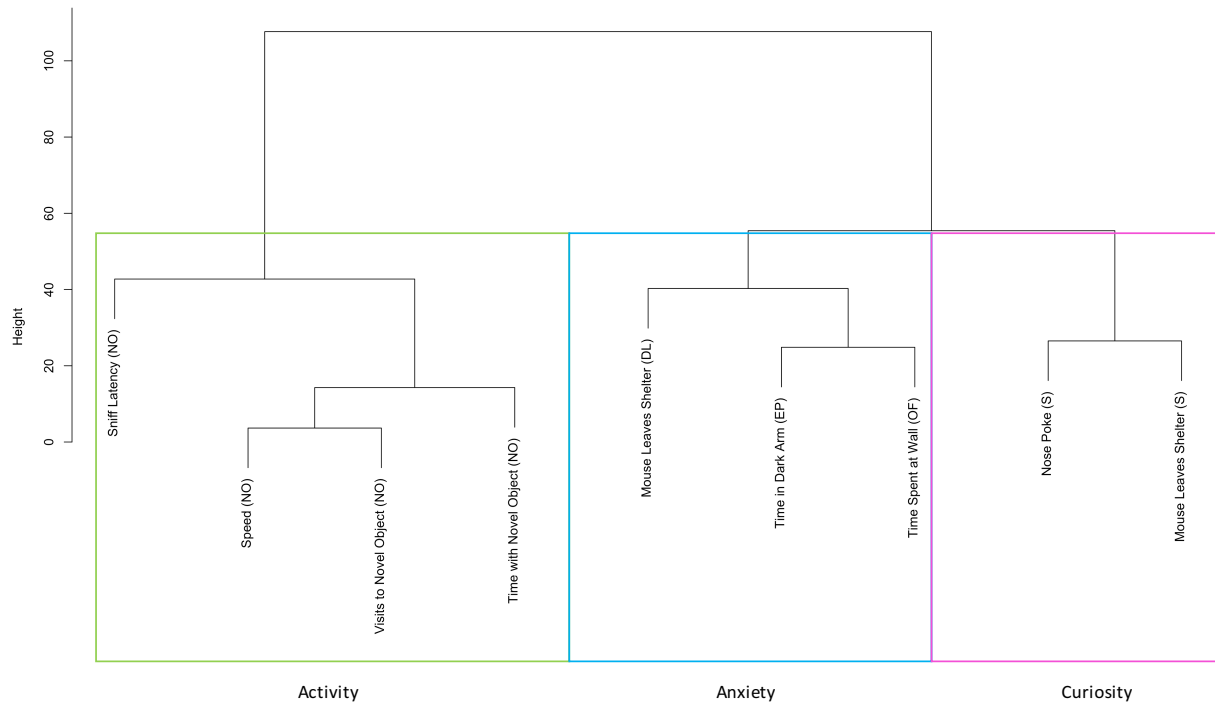


Figure D.13: Correlation plot for repeatable measurements. Correlation plot for the nine measurements from four different behavioural tests, which were repeatable across the course of the experiment. Novel Object (NO), Dark/Light box (DL), Startle Test (S), Novel Object Test (NO), Elevated Plus Maze (EP) and Open Field Test (OF) (n=33).



Agglomerative Coefficient=0.79

Figure D.14: Hierarchical clustering of repeatable measurements. Height of branches indicates similarity between measurements. Measurements could be pooled into 2 or three main categories, later classified as curiosity, anxiety and activity based on the types of measurements in the clusters. Novel Object (NO), Dark/Light box (DL), Startle Test (S), Novel Object Test (NO), Elevated Plus Maze (EP) and Open Field Test (OF) (n=33).

Behavioural Scores

Behavioural scores were calculated based on the clusters in Figure D.14. Principle component analysis was completed using these measurements and behavioural scores were generated by using the first principle component for each cluster. In this way, many variables can be combined to make one measure for activity (Figure D.15.A), anxiety (Figure D.15.B) and curiosity (Figure D.15.C). There were significant differences in the overall scores between populations for activity (Figure D.16.A, $t = 3.3753$, $df = 24.185$, $p\text{-value} = 0.002487$) and curiosity was approaching significance (Figure D.16.C, $t = 1.8489$, $df = 30.788$, $p\text{-value} = 0.0741$) but not for the anxiety score (Figure D.16.B, $t = -0.17796$, $df = 30.611$, $p\text{-value} = 0.8599$).

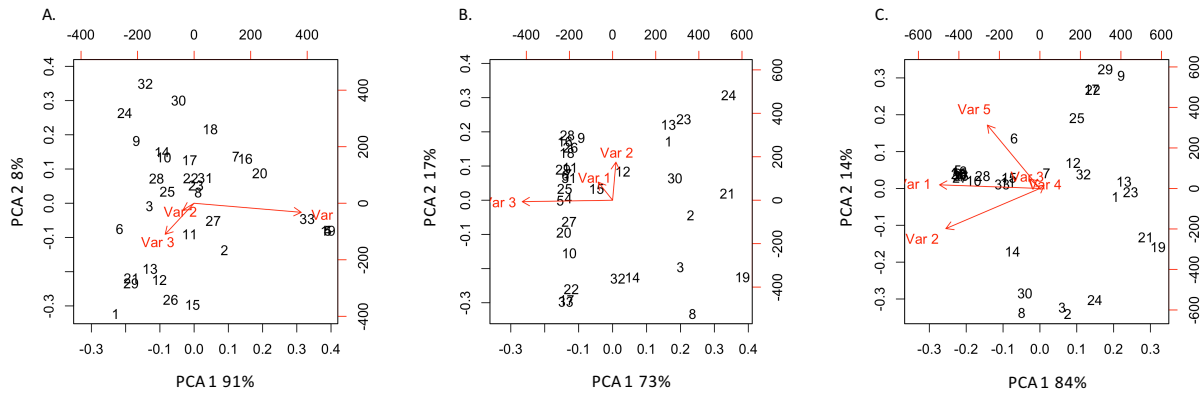


Figure D.15: Behavioural scores. A. PCA 1 for activity related measurements represents 91% (Var1 = Speed (NO), Var2= Sniff Latency (NO), Var3=Visits to Novel Object (NO), Var4=Time Spent with Novel Object (NO), B. PCA 1 for anxiety represented 73% of the variance (Var1= Time Spent at Wall (OF), Var2=Time in Dark Arm (EP), Var3=Mouse Leaves Shelter (DL), C. PCA 1 for curiosity represents 84% of the variance (Var1= Nose Poke (S), Var2 = Mouse emerges from shelter after startle (S) (n=33).

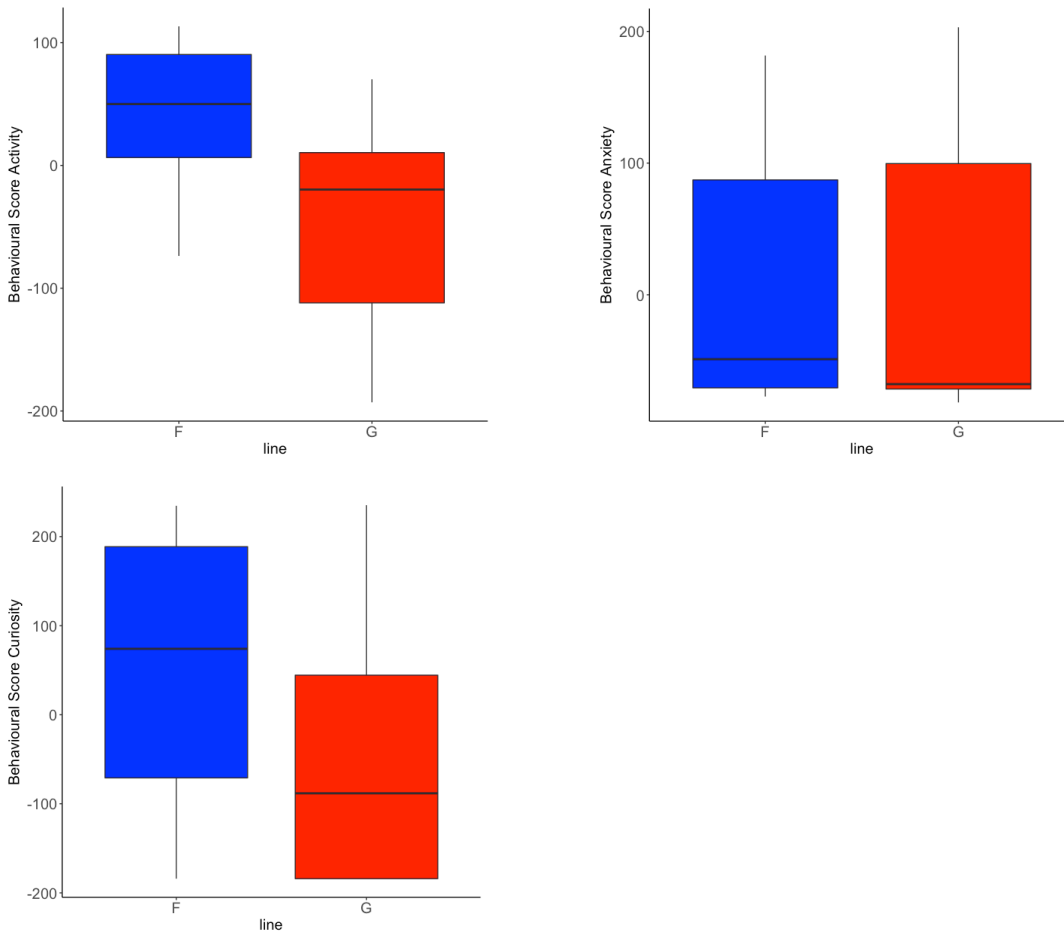


Figure D.16: Creating behavioural scores by population. A. Activity score (PCA1) by population origin of mice, B. Anxiety score (PCA1) by population and C. Curiosity score (PCA1) by population origin of mice for n=33 mice.

Quick Tests

Results from the quick tests (jumps, bites, squeaks and tail jerks) were score based on presence or absence of behaviour, rather than count. Generalized linear models from a binomial family were used to score data, using enclosure origin and monitoring number as random effects and using sex and status as fixed effects. Starting with jumps, male mice were slightly more likely to jump than female mice (Table D.9). There was no significant effect of sex or status on bites (Table D.10) or squeaks (Table D.11), but there was a slight difference between males and females in tail jerk (Table D.12).

Table D.9 Generalized linear models for jumps in quick tests. Generalized linear model for jumps using sex and status as fixed effects and monitoring as a random effect (enclosure was dropped because variance was equal to 0). Simplest model was accepted based on AIC 387.7).

Fixed Effects	Estimate	SE	z-value	P-value
Intercept	-2.72	0.33	-8.23	<0.001
Status	0.411	0.32	1.26	0.20
Sex	0.632	0.27	2.29	0.02
Random Effects	Variance	SD		
Monitoring	0.04	0.21		

Table D.10 Generalized linear models for bites in quick tests. Generalized linear model from a binomial family for bites using sex and status as fixed effects and monitoring and enclosure were kept as a random effects. Simplest model was accepted based on AIC 619.1).

Fixed Effects	Estimate	SE	z-value	P-value
Intercept	-1.31	0.43	-2.97	0.002
Status	-0.19	0.22	-0.87	0.38
Sex	0.17	0.20	0.87	0.38
Random Effects	Variance	SD		
Monitoring	0.04	0.21		
Enclosure	0.39	0.63		

Table D.11 Generalized linear models for squeaks in quick tests. Generalized linear model from a binomial family for squeaks using sex and status as fixed effects and monitoring and enclosure were kept as random effects. Simplest model was accepted based on AIC 393.9).

Fixed Effects	Estimate	SE	z-value	P-value
Intercept	-1.91	0.52	-3.64	<0.001
Status	-0.04	0.31	-0.14	0.88
Sex	0.26	0.27	0.96	0.33
Random Effects	Variance	SD		
Monitoring	0.53	0.72		
Enclosure	0.33	0.58		

Table D.12 Generalized linear models for tail jerks in quick tests. Generalized linear model from a binomial family for tail jerks using sex and status as fixed effects and monitoring and enclosure were kept as a random effects. Simplest model was accepted based on AIC 390.8).

Fixed Effects	Estimate	SE	z-value	P-value
Intercept	-2.81	0.49	-5.69	<0.001
Status	0.45	0.37	1.24	0.21
Sex	0.52	0.27	1.91	0.06
Random Effects	Variance	SD		
Monitoring	0.31	0.55		
Enclosure	0.15	0.38		

Individual Plasticity

Behavioural Reaction Norms

There was some variation between all individual measurements over the course of the experiment. These measurements could just be “noise” or a kind of plasticity in behaviours caused by experience (habituation) or seasonal changes. Behavioural reaction norms (BRN) were calculated for each individual between the first two measurements taken during the experiment showing the general variation between measurement and individual. To control for any effect that may have come from missing individuals as time goes on, only individuals with both measurements were included in the models. Unfortunately, this reduced the power of some of the tests (n=33 individuals), so some measurements had to be dropped because the models could not converge. These models give us some idea of individual consistency, changes in variance between monitoring one and two, and show the differences between the MC/F and CB/G populations. LMM models were made and random slopes were calculated for individuals and monitoring. To visualize plasticity of all measured behavioural traits, the random effects were plotted.

For the dark/light tests (n=33) the variance for individual x monitoring random effects for the first nose poke (Figure D.17.A, Table D.13), for the first time the mouse leaves the shelter (Figure D.17.B, Table D.14) and for the startle test, the first time the mouse pokes it’s nose out of the shelter post startle (Figure D.17.C, Table D.15) and the first time the mouse leaves the shelter post startle (Figure D.17.D, Table D.16). LMM models are summarized below in Table D.11 to Table D.16. In general, there was no visible difference between the variance between populations and there was no consistent pattern

(increase, decrease or interaction) between the first and second measurement. Variance did not show a consistent pattern between monitoring 1 and 2.

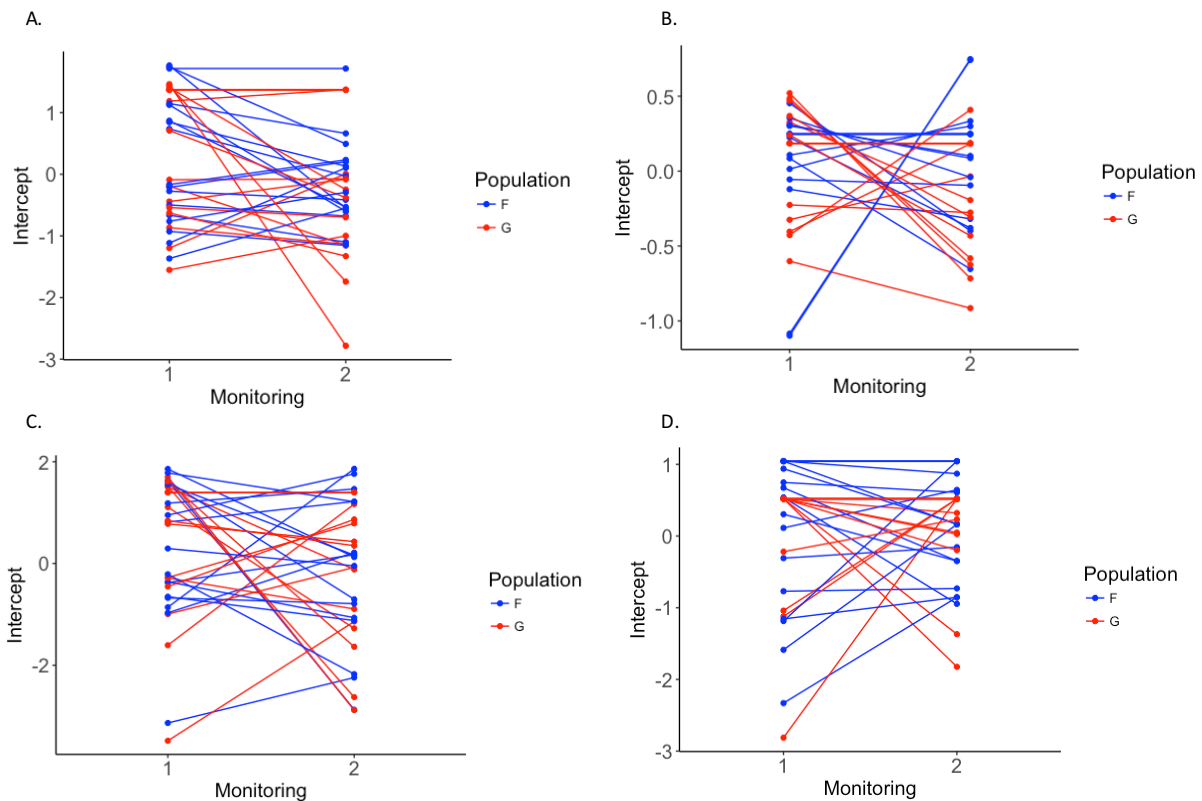


Figure D.17: Behavioural reaction norms for Dark/Light Box and Startle Test. A. Nose Poke in Dark/light Box, B. the first time the mouse emerges from the shelter during the dark light box, C. the first time the nose pokes out of the shelter after startle and D. the first time the mouse emerges from the shelter post startle (n=33).

Table D.13 Generalized linear model with random effects. This model was used to generate the random effects and slopes used in Figure D.11.A. for nose poke in the Dark/light box. Random effects included an interaction between the monitoring (1 or 2) and the ID of the individual mouse.

Fixed Effects	Estimate	SE	z-value	P-value
Intercept	4.2685	0.1645	25.94	<0.001
Population	-0.3628	-1.575	-1.575	0.115
Random Effects	Variance	SD		
Monitoring/ID	1.21	1.10		
ID	0.027	0.166		

Table D.14 Generalized linear model with random effects. This model was used to generate the random effects and slopes used in Figure D.11.B. for mouse emerges from the shelter in the Dark/light box. Random effects included an interaction between the monitoring (1 or 2) and the ID of the individual mouse.

Fixed Effects	Estimate	SE	z-value	P-value
Intercept	5.294	0.1136	49.61	<0.001
Population	-0.137	0.159	-0.86	0.388
Random Effects	Variance	SD		
Monitoring/ID	0.279	0.528		
ID	0.166	0.409		

Table D.15 Generalized linear model with random effects. This model was used to generate the random effects and slopes used in Figure D.11.C. for the time it took for the mouse to look out of the shelter in the startle test. Random effects included an interaction between the monitoring (1 or 2) and the ID of the individual mouse.

Fixed Effects	Estimate	SE	z-value	P-value
Intercept	4.089	0.25	15.84	<0.001
Population	-0.398	0.36	-1.10	0.27
Random Effects	Variance	SD		
Monitoring/ID	2.70	1.645		
ID	0.206	0.454		

Table D.16 Generalized linear model with random effects. This model was used to generate the random effects and slopes used in Figure D.11.D. for the time it took for the mouse to emerge from the shelter in the startle test. Random effects included an interaction between the monitoring (1 or 2) and the ID of the individual mouse.

Fixed Effects	Estimate	SE	z-value	P-value
Intercept	5.183	0.14	35.00	<0.001
Population	-0.526	0.20	-2.53	0.0112
Random Effects	Variance	SD		
Monitoring/ID	1.04	1.02		
ID	<0.001	<0.001		

For the novel object test, not all measurements could be fitted with a model. For sniff latency (Figure D.18.A) and the amount of time the mouse spent with the novel object (Figure D.18.B) this was similar. Both MC/F and CB/G individuals varied between the measurements and there no consistent negative or positive trend. This was the same for the other measurements, the open field test time at wall (Figure D.19.A) and the distance traveled (Figure D.19.B) and the elevated plus maze (Figure D.20.A). In the end, it seems that there was no overall effect on variance due to time on the individual measurements (at least not one that was individually consistent).

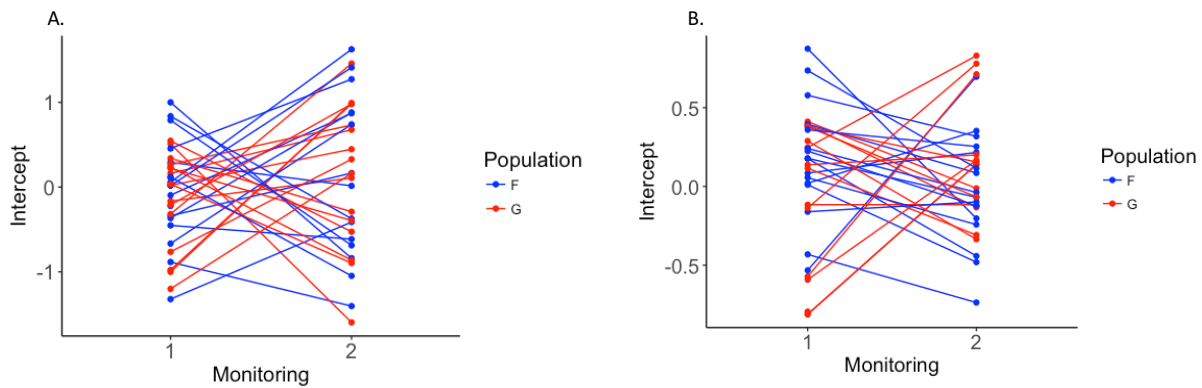


Figure D.18: Behavioural reaction norms for Novel Object Test. A. Sniff latency in the novel object test, B. the time spent with the novel object (n=33).

Table D.17 Generalized linear model with random effects. This model was used to generate the random effects and slopes used in Figure D.18.A. for the sniff latency in the novel object test. Random effects included an interaction between the monitoring (1 or 2) and the ID of the individual mouse.

Fixed Effects	Estimate	SE	z-value	P-value
Intercept	5.183	0.14	35.00	<0.001
Population	-0.526	0.20	-2.53	0.0112
Random Effects	Variance	SD		
Monitoring/ID	1.04	1.02		
ID	<0.001	<0.001		

Table D.18 Generalized linear model with random effects. This model was used to generate the random effects and slopes used in Figure D.18.B. for the time spent with the novel object. Random effects included an interaction between the monitoring (1 or 2) and the ID of the individual mouse.

Fixed Effects	Estimate	SE	z-value	P-value
Intercept	5.183	0.14	35.00	<0.001
Population	-0.526	0.20	-2.53	0.0112
Random Effects	Variance	SD		
Monitoring/ID	1.04	1.02		
ID	<0.001	<0.001		

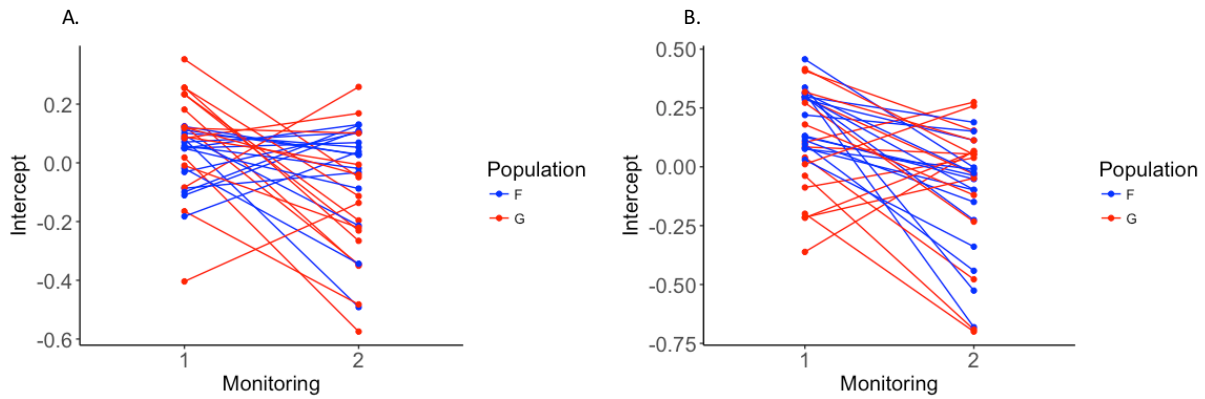


Figure D.19: Behavioural reaction norms for Open Field Test. A. Time spent at the wall in open field test B. distance traveled during open field test (n=33).

Table D.19 Generalized linear model with random effects. This model was used to generate the random effects and slopes used in Figure D.19.A. for the time spent at the wall during the open field test. Random effects included an interaction between the monitoring (1 or 2) and the ID of the individual mouse.

Fixed Effects	Estimate	SE	z-value	P-value
Intercept	5.18	0.04	120.03	<0.001
Population	0.20	0.06	3.44	0.0005
Random Effects	Variance	SD		
Monitoring/ID	0.051	0.225		
ID	0.013	0.114		

Table D.20 Generalized linear model with random effects. This model was used to generate the random effects and slopes used in Figure D.19.B. for the distance traveled in the open field test. Random effects included an interaction between the monitoring (1 or 2) and the ID of the individual mouse.

Fixed Effects	Estimate	SE	z-value	P-value
Intercept	7.81	0.05	153.16	<0.001
Population	0.05	0.07	0.75	0.454
Random Effects	Variance	SD		
Monitoring/ID	0.09	0.30		
ID	0.01	0.10		

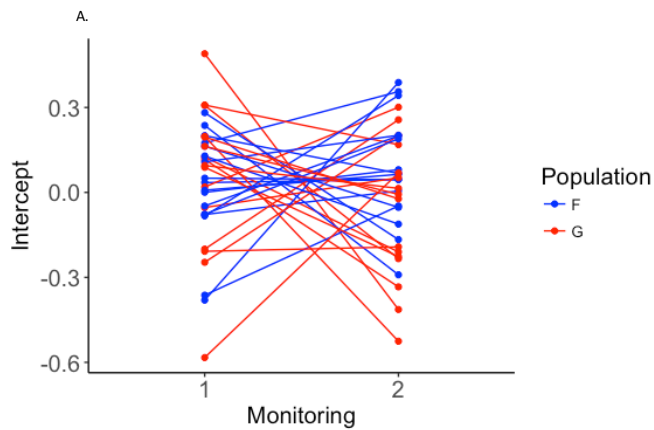


Figure D.20: Behavioural reaction norms for Elevated Plus Maze. A. Time spent in the dark arm of the elevated plus maze (n=33).

Table D.21 Generalized linear model with random effects. This model was used to generate the random effects and slopes used in Figure D.20 for the time spent in the dark arm in the elevated maze. Random effects included an interaction between the monitoring (1 or 2) and the ID of the individual mouse.

Fixed Effects	Estimate	SE	z-value	P-value
Intercept	5.02	0.04	109.4	<0.001
Population	0.08	0.06	1.3	0.194
Random Effects	Variance	SD		
Monitoring/ID	0.09	0.30		
ID	0	0		

E. Fitness in house mice

Fitness of house mice could be influenced by personality type and the way the mice use the space in the enclosures. Behavioural scores from section D were combined with the absolute number of offspring produced by individuals in order to understand the role this may play in influencing fitness. Further, territory use scores from section C were also compared to the absolute number of offspring produced. In this way, it was hoped we could determine the role behaviour may play in influencing fitness of wild house mice in our enclosures.

Personality and fitness

To identify a potential relationship between personality traits and fitness (i.e. the number of offspring), the behavioural scores measured for activity, anxiety and curiosity were compared to the absolute number of offspring produced. As personality tests have only been carried out for founder males and a

subset of G1 males (n=5), the initial analysis was restricted to G0 males, due to missing data from the G1 group. The 5 G1 male mice 2 mixed and 3 pure status produced between 2 and 0 offspring. Generalized linear mixed effects model were used, with population and activity score, anxiety score and curiosity score as fixed and enclosure as random factor (Table E.1). CB/G mice produced fewer offspring overall than MC/F mice (Table E.1). CB/G mice had lower activity scores overall compared to MC/F mice. There was an interaction between the activity score and population, where MC/F mice which were more active produced more offspring, while CB/G mice which were less active produced more offspring (Figure E.1.A). MC/F mice with lower anxiety scores (less anxious mice) produced more offspring during the experiment (Table E.1), while CB/G mice with higher scores produced more offspring (Figure E.1.B). Finally, mice which were more curious produced more offspring overall, with no interaction between population (Figure.E.1.C, Table E.1).

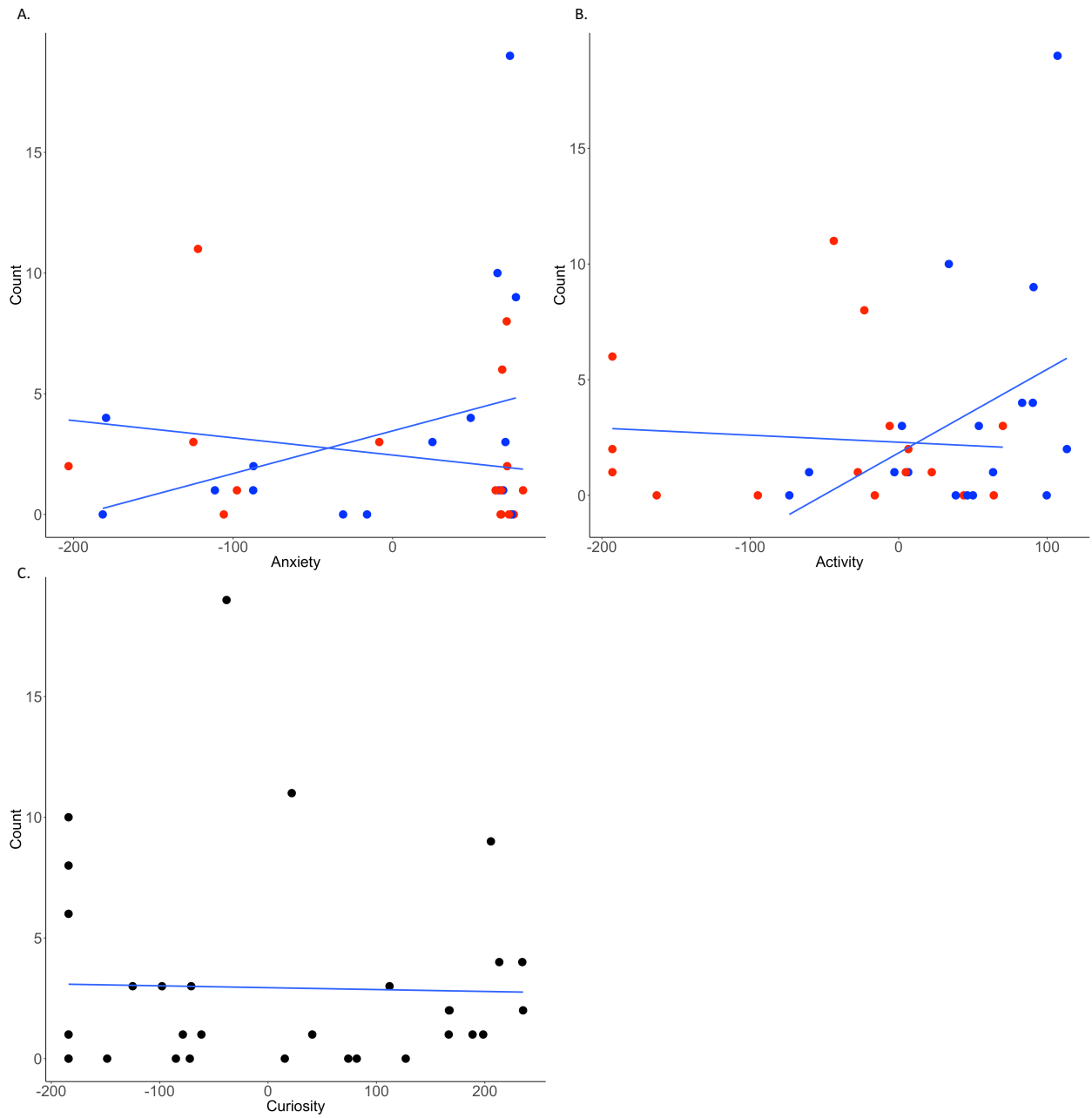


Figure E.1: A. Behavioural score for activity by the absolute number of offspring they produced for founder male mice (n=33). B. Behavioural score for anxiety by the absolute number of offspring they produced for founder male mice by population origin (n=33) and C. the behavioural score for curiosity by absolute offspring count (n=33).

Table E.1: Fixed effects of behavioural scores in relation to total number of offspring. A generalized mixed model with poisson distribution and log link was used with the absolute offspring count as response variable and personality scores and population background as fixed effects. Enclosure was included as random effect. In the simplest model (AIC 170.4), the score for curiosity was dropped as a fixed effect to gain the best fit. Enclosure was kept as a random effect because variance was greater than 0 (n=33).

Fixed Effects	Estimate	SE	z-value	P-value (Wald)
Intercept	-0.19	0.39	-0.47	
Anxiety Score	8.06	2.16	3.77	<0.001
Population	0.96	0.41	2.38	0.019
Activity Score	20.65	4.37	4.73	<0.001
Curiosity Score	-1.97	0.88	-2.22	0.025
Activity: Population	-12.84	2.69	-4.76	<0.001
Anxiety: Population	-21.20	4.86	-4.35	<0.001
Random Effects	Variance			
Enclosure	0.09			

Three factors are influencing the total number of offspring A. Anxiety seems to influence the total number of offspring such that less anxious mice are more successful B. activity and population background, which show significant effects on their own but also being significant in interaction. More active mice produced more offspring, and MC/F mice tended to be more active than CB/G mice.

Territorial behaviour and fitness

Fitness of Founder Male Mice

The scores for territory use (AA, HRS and SI) of founder male mice were used as variables to explain the absolute offspring count of mice, using enclosure origin as a random effect in the model generalized mixed effects model (Table E.2.A). HRS and AA, which were inversely related to one another, showed the

same pattern in these models. Mice with longer HRS and higher AA produced the most offspring (Figure E.2.A and Figure E.2.B respectively). In the case of SI, it seems that founder male mice which produced the most offspring were mice with intermediate scores (Figure E.2.C).

Table E.2 Generalized mixed model for a poisson distribution with log link for founder male mice (n=33). The absolute offspring count using territory use scores (AA, HRS, and SI) and population as fixed effects and enclosure as a random effect. The full model did best explain the data (AIC 139.3) Enclosure was kept as a random effect because variance was greater than 0 (n=33, MC=17, CB=16).

Fixed Effects	Estimate	SE	z-value	P-value (Wald)
Intercept	2.9932	0.9024	3.317	0.0009
HRS	-0.4408	0.1349	-3.268	0.0010
SI	-3.1055	1.0253	-3.029	0.0025
AA	-0.1276	0.0307	4.150	<0.001
Population	-0.3151	0.2209	-1.426	0.1539
Random Effects	Variance			
Enclosure	0.30			

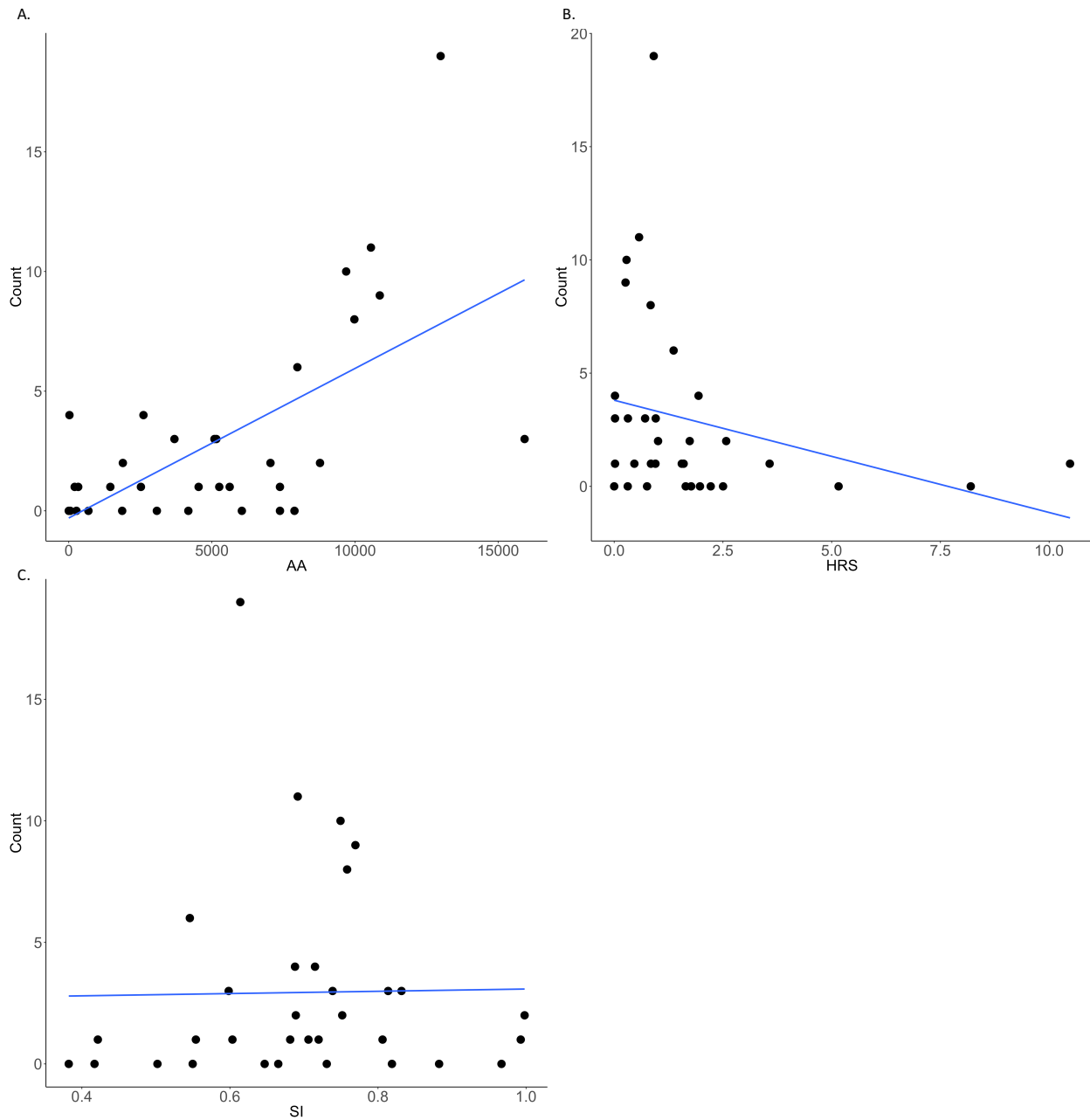


Figure E.2: A. AA score B. HRS and C. SI score by the absolute number of offspring they produced for founder male mice by population origin (n=33).

Fitness for all mice

A generalized mixed model was used to compare the territory scores of all individuals which had all three scores (mice which less than 5 registered location changes could not be scored). This left 188 individuals, 92 males and 96 females, whose absolute offspring numbers could be compared to territorial scores. In addition to the factors included in the model for only founder males, we added sex as potential influencing factor. The model showed that like the founder male mice, there was a relationship between AA and SI, although not with HRS alone (Table E.3). There was an interaction between AA and sex (Figure E.3.A) and HRS (Figure E.3.B), showing that territory size and activity affected the absolute offspring count in different ways; namely, that higher activity levels (AA) and lower territory size (HRS) were associated with higher fitness. Like the founder males, mice with intermediate SI scores tended to have more offspring (Figure E.3.C). Finally, there was an overall difference in offspring produced by an individual between males and females (Table E.3).

Table E.3 Generalized mixed model for a poisson distribution with log link. The absolute offspring count for all mice using territory use scores (AA, HRS, and SI) and sex as fixed effects and enclosure as a random effect. In the simplest model (AIC 779.6), all variables had to be kept. Enclosure was kept as a random effect because variance was greater than 0.

Fixed Effects	Estimate	SE	z-value	P-value (Wald)
Intercept	1.3723	0.4541	3.022	<0.001
AA	0.0711	0.0142	4.988	<0.001
Sex	0.3048	0.2764	1.103	<0.001
SI	-2.4916	0.5955	-4.184	<0.001
HRS	0.0359	0.0540	0.664	0.32979
AA:Sex	0.0465	0.0217	2.136	0.03270
HRS:Sex	-0.1417	0.0702	-2.020	0.04341
Random Effects	Variance			
Enclosure	0.024			

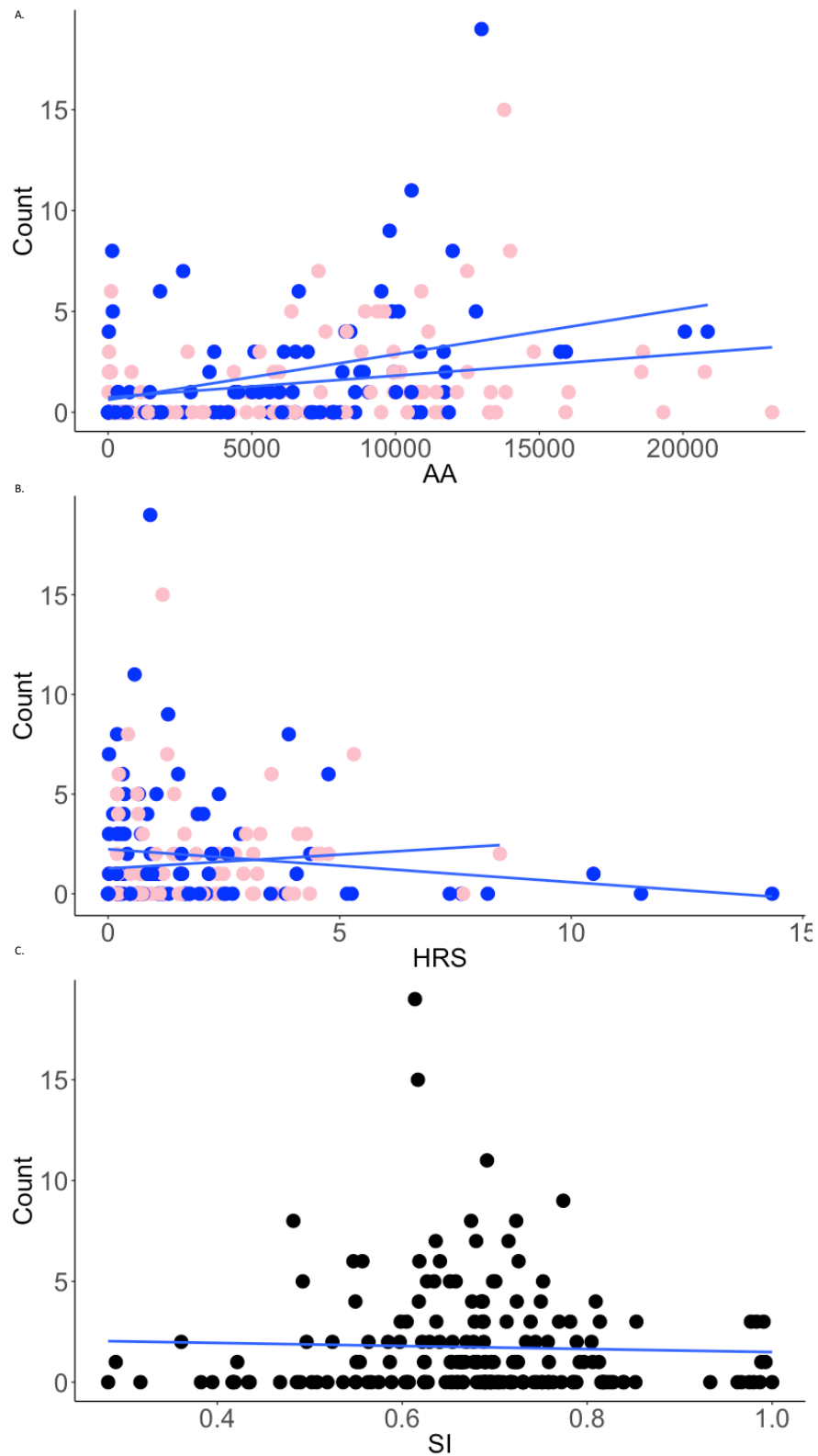


Figure E.3: A. AA score B. HRS by the absolute number of offspring they produced for all mice with all three territory scores by sex and C. SI for all mice with all three territory scores (n=188).

4. Discussion

This thesis addresses the question about personality and territory use and its influence on fitness in wild mice. We performed a semi-natural enclosure experiment with mice from two populations (France and Germany). Further we conducted regular personality tests on male founder mice and parentage analysis for all mice. Besides studying personality and territory use, this setup allowed to revisit the mate choice experiments carried out by Montero *et al.* (2013) and take another look at the mating pattern which was discovered, in which G1 mice chose mates with a similar population origin as their father (Montero *et al.* 2013).

Population development

All three enclosures started at the same time (mid-April 2015) and ran for 10 months. In the beginning of the experiment, there was some delay in breeding and on average the first new mice were born one and a half months after the experiment began. After this point, the population in the enclosures increased slowly, but steadily until the end of the experiment. At the end of the experiment the population density in our enclosures was between 7.7 and 8.4 mice per square metre. The time it took for the first litters to be born was similar to that found in previous enclosure experiments (Montero *et al.* 2013). Our population growth was much slower; however, and despite lengthening the time of the experiment, the population density was never as high as that from the previous enclosure experiments (between 11.2 and 12.9 mice per square metre after 196 days Montero *et al.* 2013). This could be due to differences in the enclosure setup. In order to create a “realistic” habitat for the mice and in an attempt to make differences in territory quality, food and water were placed at the front of the enclosure, instead of distributing them throughout as Montero *et al.* (2013) did. Noyes *et al.* (1982) studied feral house mouse populations in outdoor enclosures in the United States. They compared populations in which resources (food) were centralized or evenly distributed, based on the hypothesis that social structure in wild populations of house mice may modulate population density through social interactions (Noyes *et al.* 1982). In their study, they measured the frequency of urine marking to establish whether male mice were dominant or subordinate, which they believed would be related to population density and spatial use. They found increased frequency of urine marking (dominant behaviour) preceded decreases in population size in both centralized and decentralized populations. Furthermore, in the case of centralized populations the most dominant (highest marking) individuals were trapped more often in proximity to resources or structures where resources would be held than subordinate individuals (Noyes

et al. 1982). They conclude that centralized resources force social interactions which may regulate population size. Such an effect could explain the lower population densities in our experiments if dominant males were controlling access to resources. Differences could also be due to the antennae used in our set up, although it is unlikely that the antenna themselves disturbed the mice, because similar antennae systems have been used very successfully to track the movements of mice in other experiments (Koenig *et al.* 2015). One other modification we made was giving each nest box only one entrance/exit, so it was possible to track the directional movement of the mice, which also has been shown to not negatively influence the population development (Koenig *et al.* 2015). Nevertheless, Montero *et al.* (2013) used nest boxes with two entrances/exits to provide a possibility to escape. Furthermore, the behavioural tests may have disturbed the enclosure populations, since mice were removed from the enclosures during monitoring events. To minimize the stress caused by this process, mothers caught in houses were kept in their homes with their pups. Unfortunately, mothers which were not caught in houses could not be kept with their offspring, and litter which were hidden in the enclosures may have been lost without maternal supervision until their mothers were returned to the enclosures a few hours later. Finally, it is also possible that the experiments were affected by seasonal changes in temperature and humidity in the enclosures, which were not present under the controlled conditions of Montero *et al.* (2013). From these results, we can only speculate about what may have reduced the population size during our enclosure experiments, but future experiments should be planned carefully and the potential stress to mice caused by additional disruptions should not be underestimated.

Sex ratios differed in the enclosures. In enclosure C, the sex ratio was even at the end of the experiment, in enclosure A it was also even, but bordering on a statistical difference and enclosure B had significantly more females than males at the end of the experiment. Since we did not check the sex ratio of pups during the experiment (to avoid disturbing nests), it is not possible to say whether the sex ratio resulted from differences in litter composition or differential survival due to competition in the enclosures. It is suggested that overcrowding (poor nutrition or increased competition) may contribute to unbalanced sex ratios in litters (Wright *et al.* 1988). In this case, poor nutrition resulted in female biased litters. This is consistent with the theory made by Trivers and Willard (1973), suggesting that mothers may bias their investment in offspring from male to female depending on environmental conditions. However, the distorted sex ratio in one of the enclosures probably is not due to restricted resources as food and water were given ad libitum with a lot of surplus. Finally, the uneven sex ratio may have been due to increased male-male interactions around centralized resources (Noyes *et al.* 1982). Male-male aggression in

commensal populations is higher around resources and may have led to increased mortality among male mice, thus skewing the sex ratio (Gray and Hurst, 1997; Frynta *et al.* 2005).

Mate Choice

One of the main goals of our experiment was to test for the mating pattern found in Montero *et al.* (2013). Similar to Montero *et al.* (2013), we found no pattern of assortative mating in the initial founder to founder crosses. Unlike Montero *et al.* (2013), in which a pattern of paternal matching was demonstrated, we did not find a significant pattern of paternal matching in the crosses between the first generation of mice born in the experiment (G1 x G1), although there appeared to be a tendency towards assortative mating with an increased number of GF/GF and GG/GG crosses. Nevertheless, unexpectedly we did not reach the same level of population density in our enclosures as discussed above, thus we might lack the statistical power to detect such a pattern. Further, a recent study by Linnenbrink and von Merten, 2017 also detected a paternal influence on its offspring's partner choice, which is based on the father's MHC. As in our study we did not take MHC or other genetic loci potentially being involved in mate choice (e.g. MUPs), into account, we can't confirm or support their findings. Nevertheless, it would be interesting to add information from our semi-natural experiment to their findings from a controlled cage experiment.

Besides comparing the population of origin of both mating partners to potentially find a certain mating pattern, we used the number of offspring and the number of successful mating events to compare different groups of mice. These groups were pure (FF and GG) vs. mixed (FG and GF) individuals, males vs. females, mice from the MC/F vs. CB/G population. Similar to Montero, 2010, we also found no significant difference between pure vs. mixed status mice in offspring count or mating count; however, some values (absolute offspring count of male pure vs. mixed G1 and mating count for female pure vs. mixed G1) are approaching significance. This pattern disappeared in the time adjusted relative scores for successful mating and offspring count.

Territory use

Our approach to study territory use was to track the presence of mice in the enclosures by using RFID technology. Mice were tagged with a RFID transponder and RFID antennae were distributed in the enclosures, especially water bottles, nest boxes and shelters. Indices for territory use were HRS,

selectivity of certain nest boxes and the activity of the mice in the enclosures. There should be some fitness advantage to maintaining a territory, since the main purpose of maintaining a territory should be to secure resources, such as nutrition or access to mates (Brown and Orians, 1970). We found that there was a moderate correlation between the average HRS of fathers and that of their offspring. Furthermore, there was a moderate correlation between the offspring HRS of HRS to the mean HRS of both parents (Figure C.8). There was also a significant, but weaker correlation between the SI score of mothers to their offspring (Figure C.9). AA scores were related to sex, status (pure vs. mixed) and the population background of parents, HRS was related to status and the population background of the mother, and SI was related to sex and the population background of parents. Enclosure did not influence overall mean AA or SI and was excluded as a random effect from models (Table C.1. and C.3); however it was a source of variance in HRS measurements (Table C.2). The identity of the father also contributed to the variance in scores of all three territorial scores (Table C.1, C.2 and C.3). Finally, we found that fitness was related to HRS, and that males with a smaller HRS produced more offspring, while females with a greater HRS were more successful (Figure E.3.B.).

Enclosure experiments have been carried out using house mice to study territory use since the 1960's (Crowcroft and Rowe, 1963). These experiments usually involved indoor enclosures (Crowcroft and Rowe, 1963; Singleton and Hay, 1983; Montero *et al.* 2013) and outdoor enclosures (Lidicker, 1976; Noyes *et al.* 1982). These studies focused on social structure, territorial behaviour and/or population density and varied in size from 375 metres (Lidicker *et al.* 1976, outdoor) squared to 16.9 metres squared (Singleton and Hay, 1983, indoor). The indoor enclosure experiments vary slightly from one another in size, number of nest boxes, etc. and those that mentioned the distribution of resources (food) and shelter, used decentralized or uniform distributions (Crowcroft and Rowe, 1963; Montero *et al.* 2013). There is no information on the distribution of nutrition in Singleton and Hay (1983). Only Noyes *et al.* (1982) studied populations with centralized nutritional resources, similar to our experiment. Similar to our experiments, none of these experiments were long-term studies. Feral populations of house mice have also been studied, to understand movement and territory use (Newsome, 1969; Selander 1970; Fitzgerald *et al.* 1981; Singleton, 1983; Krebs *et al.* 1995; Chambers *et al.* 2000). Large field studies like Krebs *et al.* (1995) and Chambers *et al.* (2000) studied wild populations in order to understand the ecological factors which contribute to mouse plagues in Australia. These studies focused on HRS (mostly lower than 1000 metres squared) and overlap in feral mouse populations using radio tracking methods in Queensland, Australia. Krebs *et al.* (1995) found that in feral populations of house mice there are large overlaps in territories during breeding periods, males have larger territories and move around more

often than females. They also found that when mice were not breeding, HRS increased dramatically (10 fold) and that mice were more likely to be nomadic (Krebs *et al.* 1995). These findings were similar to those of Chambers *et al.* (2000) in Victoria, Australia. They found sex specific differences in spatial use between males and females, females being relatively consistent over time, showing very little overlap with other females, while males overlap more in non-breeding seasons, showing behavioural plasticity in social/territorial behaviour (Chambers *et al.* 2000). This is in contrast to Selander (1970) and Fitzgerald *et al.* (1981) which reported stable territory use regardless of the time of year (Chambers *et al.* 2000). Commensal mice, unlike feral populations, have much smaller HRSs and live at much higher densities which leads to more hierarchical behaviour, territorial behaviour and increased social interactions (Gray and Hurst, 1997; Frynta *et al.* 2005).

In our experiments we found sex specific differences in the mean AA scores and SI scores. Plots of AA scores by monitoring periods show that female scores were usually higher. Mean male SI scores were higher overall than females, and increased slowly over the course of time. It is important to remember that the AA score is related to antenna use and not necessarily to activity. AA scores are negatively correlated with HRS, AA is higher when HRS is low. We found much lower HRS than Krebs *et al.* (1995) and Chambers *et al.* (2000), which is the expectation for commensal house mice in barns or semi-natural enclosures (Crowcroft, 1967; Frynta *et al.* 2005; Wolff *et al.* 2009). Unlike Krebs *et al.* (1995) and Chambers *et al.* (2000), we did not find sex specific differences in mean HRS overall, nor did we find a switch in male HRS over time. This makes sense, since the switch corresponded to the breeding season, which does not affect commensal mice, which can breed year round. Territory scores for AA, SI and HRS were also influenced by the status of mice (pure or mixed). Overall differences between pure and mixed individuals may be related to the age of individuals, founders were included in the pure individuals and mixed individuals are all younger adults. Further analysis with time series is needed to separate the factors influencing these scores. HRS measured by monitoring period shows that mice from a mixed background started off with much smaller HRS than pure, but over time the average scores of mixed background mice increase. Again, this could be due to differences in age between the two groups at the beginning of the experiment and it is possible that it took time for them to establish territories within the hierarchy (Wolff, 2009). Finally, the population background of mothers influenced all territorial measurements, while the population background of the father also influenced AA. SI and HRS scores were also strongly related to the identity of their father. Meaning that the scores of an individual father were had similar scores to one another.

Unlike suggested in Wolff (2009), young mice did not displace the original territory holders (see Results section, and supplemental data for HRS plots). There was also a relationship between the HRS of G1 mice and the mean score of their parents. Comparing the HRS plots of G1 mice to their parents, it seems both male and female mice generally stayed in the home range of their father, therefore would have a similar HRS to their father and mother. Montero *et al.* (2013) also showed a pattern of offspring staying within “family” territory and familiarity playing an important role in mate choice. Immelmann (1975) considered habitat choice being an example of ecological imprinting influenced by both heritability and early life experience, citing examples in deer mice (*Peromyscus maniculatus bairdi*) (Wecker, 1963) and chipping sparrow (*Spizella passerine*) (Klopfer, 1963). Studies of HRS differences in other rodents, namely striped mouse, found that differences in territory size were related to environmental conditions and fluctuated over time and contrary to expectation, retained larger territories than were necessary to acquire resources (Schradin *et al.* 2010). Differences in territory use in our experiments may be due to social structure determined by dominant, subordinate and peripheral males and females (Singleton and Hay, 1983). It makes sense that mice which are more dominant should be able to maintain a territory, and benefit from the resources therein, while subordinate mice might be forced to adopt a different strategy. It is also possible that individuals may have different strategies for territory use, similar to those found in striped mice, where male mice adopt alternate reproductive strategies; the first, being the main breeder in a group of females; second, staying within a their original family group as a subordinate; or third, roaming (Rimbach *et al.* 2017). Finally, it is possible that the differences we observe may be influenced by the antenna system we used to track the mice, since antennae were primarily on shelters, mice which appear to have no consistent territory may have had one, they simply didn't use nest boxes. This seems unlikely to be the case, however, since the nest boxes should be the preferred territories (protection of potential predators and conspecifics). Further studies are required to understand the true nature of differences in territory use between individuals.

Personality and behavioural plasticity in wild house mice

In order to investigate if personality in wild house mice exists and to identify potential personality traits, we used a number of behavioural tests to measure personality in male founder mice. In each test, multiple measurements were taken and tested for repeatability, which is a prerequisite to the definition of personality (Gosling, 2008). Most of the behavioural measurements taken in the longer tests were repeatable across the experiment, and the scores were consistent with the repeatability scores expected

from other experiments using wild rodents (Koolhaas *et al.* 1999; Montiglio *et al.* 2012; Herde and Eccard 2013; Schuster *et al.* 2017). Test mice habituated to some experiments (dark/light box test, novel object test and open field test), which is consistent with previous experiments and attributed to a kind of plasticity, due to “a reduction in novelty” (David *et al.* 2004; Martin and Reale, 2008; Dingemanse *et al.* 2010; Herde and Eccard 2013). Looking only at individuals which had a first and second measurement for each behavioural trait measured, we were not able to find a consistent effect of context (monitoring one and two) on individual behavioural scores (based on Klueen and Brommer, 2013), rather some individuals had higher scores and others lower in the second monitoring, but there was no overall pattern related to monitoring. Rank order was also not consistent between the first and second monitoring score. Despite small differences in the scores over times, the scores for most of the measured behaviours were repeatable throughout the course of the experiment. Repeatable measurements were then clustered based on correlations to form patterns of related behaviours known as personality traits. In our case, we named these clusters activity, anxiety-like behaviour and curiosity/risk taking, although these scores could be related to activity, exploration and boldness measurements from studies in other rodents (Herde and Eccard, 2013; Schuster *et al.* 2017).

Differences in behaviour between two recently diverged populations

The two populations used in this study diverged ~3000 years ago. Nevertheless they are already well differentiated and they show differences in behaviour, e.g. ultrasonic vocalisation (von Merten *et al.* 2014), territoriality and space use (Reinhardt 2015). Furthermore, these populations are known to be genetically distinct (Ihle *et al.* 2006, Teschke *et al.* 2008, Staubach *et al.* 2012) and also differ in gene expression patterns (Bryk *et al.* 2013; Bryk and Tautz, 2014, Lorenc *et al.* 2014).

Our studies suggest that the populations used also differ in some behavioural measures. In fact, mice originating from the CB/G and MC/F population varied in behaviours related to activity and curiosity, not so with anxiety. Behavioural differences between populations are not uncommon in wild populations. Personality traits between populations of *Rana temporaria* differ between island and main land populations (Brodin *et al.* 2013). One theory is that local differences in predation risk may influence behavioural traits in different locations (Magurran, 1986; Bell, 2005; Dingemanse *et al.* 2007). Another could be that there is a founder effect resulting from expanding population or isolation (Brodin *et al.* 2013). Differences in personality have also been found between invasive *Rhinella marina* (cane toads) as

the population spreads across Queensland, Australia, where bolder individuals tend to be found at the front of the expansion (Gonzalez-Bernal *et al.* 2014).

Despite overall mean behavioural differences between the populations, both populations exhibited a full range of personality types. This is according to expectations, since personality traits are believed to be under fluctuating selection, maintaining a diversity of behavioural types within populations (Dingemanse *et al.* 2004; Le Coeur *et al.* 2015; Nicolaus *et al.* 2016). This means that the direction in which selection acts changes in relatively short periods of time (from season to season, or year to year). In this way, variation in personality types is maintained within populations. This could be a result of differences in population density (Nicolaus *et al.* 2016) or differences in resources availability, for example food availability (Dingemanse *et al.* 2004; Le Coeur *et al.* 2015). Examples of the latter include the case of Dingemanse *et al.* (2004), selection pressure was related to mast seeding in beech trees in *Parus major*, in Le Coeur *et al.* (2015), the availability of acorns (*Tamias sibiricus*). Dingemanse *et al.* (2004), write that relaxed selection pressure as a result of a rich food environment in the winter leads to increased competition for space in the spring, due to increased survival from the previous season. Dingemanse *et al.* (2004) refer to the trait of exploration speed (fast vs slow) in novel environments and suggest that while fast explorers may do better than slow explorers in times when food is scarce, the characteristic may be maladaptive in times when food is plentiful due to fighting or predation (Sih *et al.* 2003; Dingemanse *et al.* 2004). Since the mice in our study were raised under the same conditions and there was no predation in the enclosures, it is unlikely that environment caused the behavioural differences. The only condition which was constantly changing (increasing) was the population density over time, even though it did not increase into an extreme level and was in the end much less compared to the population density found by Montero *et al.* (2013). If the differences are genetic, it is possible that there are historic differences in the conditions on farms in France and Germany, which may have conferred advantages to different ranges of personality traits. It is also possible that differences arose due to random forces like drift at the front of the population which is expanding through gene surfing (Hallatschek and Nelson, 2007) or the result of selective sweeps which differentiate the MC/F and CB/G population (Staubach *et al.* 2012). In a parallel study a copy number variable locus was identified that could be causally involved in the generation of variable personalities (Maryam Keshavarz - thesis in preparation), although this does not explain per se why the two populations differ in this respect.

Personality and territory use and their influence in house mouse fitness

We found that mice which were more active in behavioural experiments produced more offspring than less active mice (Figure E.1). We also found that there was an interaction between the activity and anxiety scores and the number of offspring produced, showing that CB/G mice produced the same number of offspring regardless of activity or anxiety level, while MC/F mice did better if they were more active and less anxious. Furthermore, we found that territory use (HRS) was also related to fitness in founder males, namely that a smaller HRS was related to higher offspring count, as was an intermediate level of selectivity (Figure E.2). The pattern was similar when looking at the entire population; males produced more offspring when they had a smaller home range (E.3.B), while for female mice this was the opposite. For both male and females, an intermediate SI score resulted in more offspring (E.3.C). AA scores were inversely related to HRS scores (discussed above) and show the opposite relationship to fitness compared to the HRS pattern (E.3.A).

Personality traits can influence individual fitness because consistent behavioural patterns should influence day to day activity which directly increase or decrease the probability of survival (predation risk, conflict with conspecifics) and the opportunity to find mates and to access resources (Ingley and Johnson, 2014). Differences in fitness related to personality types have been measured and seem to be dependent on environmental conditions. Most studies focusing on personality traits and fitness have found that different personality types do better on some conditions and worse in others. Environmental conditions such as population density (Nicolaus *et al.* 2016) or food availability (Bergeron *et al.* 2013; Le Coeur *et al.* 2015) influence the fitness of individuals with different personality types. These conditions are variable and change from season to season and year to year. In our experiments food availability was consistent over time, but the population was constantly increasing. Nicolaus *et al.* (2016) and Le Coeur *et al.* (2015) found that reproductive success of different personality types was dependant of the environment and fluctuated over time. Bergeron *et al.* (2013), found a pattern of disruptive selection, where individuals with more extreme behaviours were more successful. Under the conditions of our experiment, more active and less anxious MC/F mice produced the most offspring, but it is possible that under different conditions, that might differ. It is interesting that in both cases, the effect of activity and anxiety on fitness was not as strong for CB/G mice. There are many other examples of how fitness can be influenced by individual behaviours within populations. In *Tamias sibiricus* the number of offspring produced reproductive success was linked to boldness, which varied from year to year based on acorn availability. It is possible to relate these scores to our results, since the measurements combined in the

activity score involve not only speed travelled but also the speed and frequency with which the mouse investigated a novel object. Therefore our results are similar in the short term, although we do not have long term experiments to test for seasonal or annual fluctuations in fitness. One crucial difference between our experiment and others is that we used mice from two allopatric populations which we now know showed mean differences in within behavioural traits. Under normal conditions, these populations would not be competing for the same resources in the same space. Studies of allopatric two populations in the wild *Phoxinus phoxinus* (Eurasian Minnow), Kortet *et al* 2015 found that the population which was subject to higher predation risk and increased parasite load was on average bolder but less active than those from the population with lower predation risk. Further, they found that predation and parasite load also influence survival and fitness (Seppälä and Jokela, 2008), which might also demonstrate an adaptive nature to these behavioural patterns based on environment (Kortet *et al.* 2015). In another example, Dubuc-Messier *et al.* (2017) looked a different populations of *Cyanistes caeruleus* (blue tits) living in different types of forest and exhibiting different life history traits (slow vs. fast pace of life). Specifically individuals from evergreen habitat have a slow pace of life (produce offspring later, smaller clutch/litter size and are expected to live longer than those with a fast pace of life, in this case those living in a deciduous forest, who are expected to do the opposite (Charmantier *et al.* 2016). The expectation is that individuals with lower adult survival should favour risky behaviour and fast exploration, increased survival prospective should favour less risky behaviour and slower exploration (Dubuc-Messier *et al.* 2017). Again, in this case, differences in exploration behaviour were found between the populations (Messier *et al.* 2017). Since these two populations do not come in contact in the field, it is difficult to determine why exactly these differences exist. Unfortunately, there is no literature about the comparison of allopatric populations concerning personality differences and fitness to compare our results directly.

In the case of measurements of territory use, very few studies that focused directly on HRS and fitness are currently published, although it is widely considered that there must be some cost or benefit to maintaining a large or small home range. Having a larger HRS could lead to more social contacts, both with potential mates and competitive conspecifics (Fisher and Lara, 1999) and may be able to compensate for a lack of resources or high population density by adjusting their HRS (Schoepf *et al.* 2015). Since commensal mice typically have small HRS (Gray and Hurst, 1997; Frynta *et al.* 2005) and food was provided *ad libitum*, but in a centralized location, measuring the proximity of the individual home ranges to the food may be interesting. The two male mice, which produced the most offspring, were from the MC/F population and lived directly adjacent to the feeding area. On the other hand, the

male CB/G mice which produced the most offspring were found at the back of the enclosure. Previous experiments comparing dyadic interactions between MC/F and CB/G male mice found that CB/G mice engaged in more territorial behaviour than MC/F mice, and mice which were more territorial in dyads used their opponents space more freely than subordinate mice (Reinhardt, 2015). Fitness of mice was not assessed in these experiments. As discussed above, there are examples of population differences in behaviour in other species (Bell, 2005; Dingemanse *et al.* 2007; Brodin *et al.* 2013; Gonzalez-Bernal *et al.* 2014), but differences in individual fitness are not mentioned.

5. Conclusions and Outlook

The presented work demonstrates that personality traits exist in wild mice and that certain behavioural traits indeed seem to influence the fitness (i.e. number of offspring). It is widely accepted that behavioural traits can vary over geographical space (Foster and Endler, 1999) and there are some studies focused on personality differences between populations (Bell, 2005; Dingemanse *et al.* 2007; Brodin *et al.* 2013), although there is still a lot to learn about the relationship between personality, geography and environmental elements (Bell, 2005). If personality is constrained by a shared mechanism, correlated traits should be difficult to break apart (constraint hypothesis) and if they are adaptive, correlated behaviours may differ between populations and it may be possible to break up relationships between behaviours (Bell, 2005). For example; Bell (2005) studied the relationship between boldness and aggression in different populations based on a study which found the two measurements to be related (Huntingford, 1976). Bell (2005) found that these related behaviours were easily uncoupled. In another case, Dingemanse *et al.* (2007) found that aggressiveness, activity and exploration behaviours were only correlated when there was a high risk of predation. These results suggest that personality may be more adaptive than constrained and that environmental context is important. Frynta *et al.* (2005) demonstrated that levels of aggression can be very different between populations, even within the same subspecies of house mouse. Discussions with other scientists and animal caretakers working with *M.m.domesticus* from populations outside Europe reveal that there are apparent differences between these populations in aggression, activity etc. We hypothesized that there could be differences in fitness between different personality types and it was surprising to find an interaction between population and behavioural scores when related to offspring count. It would be helpful in the future to study the fitness and personality scores of individuals from these populations separately, instead of combined in the same enclosures.

Furthermore, by using RFID technique, we gained first insight into natural territorial behaviour of house mice, specifically how territory is used and inherited to offspring. Offspring tended to stay in the territory of their parents, but what would have happened if the population density was higher? Montero (2010) reported also the existence of family clans, some of which were changing to new sites together. In our case, family clans stayed within the same territory, even young male mice, which one might have expected would try to disperse. Hence, an experimental time even longer than 10 months, i.e. following more than two generations, would be needed to enlighten a natural population development. Future studies should focus on the nature of the mice with apparently small or large territories. Mice with the smallest HRS seemed to have the most offspring, which is advantageous and it has to be asked if these mice with small territories are more dominant or social than mice which appear to roam. Further analysis of territory use and territory formation in a time series manner is needed to determine if there are overall differences between populations, males and females and/or parents and offspring. Finally, it would be interesting to use the antenna data to try to profile individuals for their circadian rhythm. In other rodent species, “chronotypes” have been measured (Labyak *et al.* 1997), to determine if there are individual difference in temporal behaviours, if this correlates with other behaviour and if it is also related to fitness in house mice.

6. References

1. Amy, M. et al. (2010). "Effects of personality on territory defence in communication networks: a playback experiment with radio-tagged great tits." Proceedings of the Royal Society B: Biological Sciences 277(1700): 3685-369
2. Archer, J. (1973). "Tests for emotionality in rats and mice-review." Animal Behaviour 21(MAY): 205-235.
3. Asaba, A. et al. (2014). "Sexual attractiveness of male chemicals and vocalizations in mice." Frontiers in Neuroscience 8(231).
4. Bates, et al. (2015). "Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software", 67(1), 1-48.
5. Benus, R. F. et al. (1991). "Heritable variation for aggression as a reflection of individual coping strategies." Experientia 47(10): 1008-1019.
6. Bell, A.M. (2005). "Behavioural differences between individuals and two populations of stickleback (*Gasterosteus aculeatus*)." J Evol Biol. 18(2):464-473.
7. Bergeron, P. et al. 2013. "Disruptive viability selection on adult exploration behaviour in eastern chipmunks." Journal of Evolutionary Biology 26(4):766-774.
8. Bergmuller, R. and M. Taborsky (2010). "Animal personality due to social niche specialisation." Trends in Ecology & Evolution 25(9): 504-511.
9. Berry, R. J. and Jakobson, M. E. (1971). "Life and death in an island population of house mouse." Experimental Gerontology 6(2): 187.
10. Berry, R. J. and Scriven, P.N. (2005). "The house mouse: a model and motor for evolutionary understanding." Biological Journal of the Linnean Society 84(3): 335-347.
11. Beynon, R.J. and Hurst, J.L. "Multiple roles of major urinary proteins in house mouse, *Mus domesticus*." Biochemical Society Transactions 31:142-146.
12. Birke, L. I. A. and Archer, J. (1983). "Some issues and problems in the study of animal exploration." van Nostrand Reinhold (UK) Co. Ltd, Wokingham, Berkshire, England.
13. Both, C. et al. (2005). "Pairs of extreme avian personalities have highest reproductive success." Journal of Animal Ecology 74(4): 667-674.
14. Bronson, F. H. (1979). "Reproductive ecology of the house mouse." Quarterly Review of Biology 54(3): 265-299.
15. Brodin, T. et al (2013). "Personality trait differences between mainland and island populations in the common frog (*Rana temporaria*)." Behavioural Ecology and Sociobiology 67: 135-143.
16. Brown, J.L. and Orians, G.H. (1970). "Spacing patterns in mobile animals." Annu. Rev. Ecol. Syst. 1:239-262.
17. Bryk, J. et al. (2013). "Early gene expression divergence between allopatric populations of the house mouse (*Mus musculus domesticus*)." Ecol. Evol. 3(3):558-568.
18. Bryk, J. and Tautz (2014). "Copy number variants and selective sweeps in natural populations of the house mouse." Frontiers in Genetics 5(153).
19. Burt, W.H. (1943). "Territory and home range concepts as applied to mammals." Journal of Mammology 24(3): 346-352.
20. Carroll, L. S. et al. (2004). "Fitness effects of a selfish gene (the Mus T complex) are revealed in an ecological context." Evolution 58(6): 1318-1328.
21. Calenge, C. (2006). The package adehabitat for the R software: a tool for the analysis of space and habitat use by animals. Ecological Modelling, 197, 516-519.
22. Cassidy, S.B. et al. (2012). "Prader-Willi syndrome." Genet. Med. 14(1):10-26.
23. Cavaille, J. (2017). "Box C/D nucleolar RNA genes and the Prader-Willi syndrome: a complex interplay." Wiley Interdiscip Rev RNA 8(4):e1417.

24. Chambers, L.K. et al. (1997). "Immunocontraception as a potential control method of wild rodent populations." Belgian Journal of Zoology 127:145-156.
25. Charmantier, A. et al. (2016). "Mediterranean blue tits as a case study of local adaptation". Evol Appl. 9(1): 135-152.
26. Crowcroft, P (1966). "Mice all over", The Whitefriars Press Ltd
27. Crowcroft, P. and Rowe, F.P. (1963). "Social organisation and territorial behaviour in the wild house mouse (*Mus musculus L.*)". Proc Zool Soc Lond 140:517-531.
28. Cucchi, T. et al. (2005). "First occurrence of the house mouse (*Mus musculus domesticus* Schwarz & Schwarz, 1943) in the Western Mediterranean: a zooarchaeological revision of subfossil occurrences." Biological Journal of the Linnean Society 84(3): 429-445.
29. Dall, S. R. X. (2004). "Behavioural biology: Fortune favours bold and shy personalities." Current Biology 14(12): R470-R472.
30. David, J.T. et al. (2004). "A neural network underlying individual differences in emotion and aggression in male golden hamsters." Neuroscience 126: 567-578.
31. Dieringer and Schloetterer (2003). "Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets." Molecular Ecology 3:167-169.
32. Ding, F. et al. (2008). "SnoRNA *Snord116* (PWcr1/MBII-85) deletion causes growth deficiency and hyperphagia in mice." PLOS One 3(3): e1709.
33. Dingemanse, N. J. et al. (2002). "Repeatability and heritability of exploratory behaviour in great tits from the wild." Animal Behaviour 64: 929-938.
34. Dingemanse, N. J., et al. (2004). "Fitness consequences of avian personalities in a fluctuating environment." Proceedings of the Royal Society B-Biological Sciences 271(1541): 847-852.
35. Dingemanse, N. J. et al. (2007). "Behavioural syndromes differ predictability between 12 populations of three spine stickleback." J Anim Ecol. 76(6):1128-1138.
36. Dingemanse, N. J. et al. (2010). "Behavioural reaction norms: animal personality meets individual plasticity." Trends in Ecology & Evolution 25(2): 81-89.
37. Drickamer, L.C. and Gillie, L.L. (1998). "Integrating proximate and ultimate causation in the study of vertebrate behaviour: methods considerations." American Zoology 38:43-58.
38. Drickamer, L.C. et al. (2003). "Free mutual mate preferences in house mice affect reproductive success and offspring performance". Animal Behaviour 65(1):105-114.
39. Dubuc-Messier, G. et al. (2017). "Environmental heterogeneity and population differences in blue tits personality traits." Behavioral Ecology 28(2): 448-459.
40. Foster, S.A. and Endler, J.A. (1999). "Geographic variation in behavior: perspectives on evolutionary mechanisms." Oxford University Press, New York.
41. Freund, J. et al. (2013). "Emergence of Individuality in Genetically Identical Mice." Science 340(6133): 756-759.
42. Freund, J. et al. (2015). "Association between exploratory activity and social individuality in genetically identical mice living in the same enriched environment." Neuroscience In Press
43. Fitzgerald et al. (1981). "Spatial organisation and ecology of a sparse population of house mice (*Mus musculus*) in a New Zealand forest." The Journal of Animal Ecology 50:489-51.
44. Frynta et al. (2005). "Aggression and commensalism in house mouse: A comparative study across Europe and the near east." Aggressive Behavior 31(3):283-293.
45. Gerschenfeld, H.K. and Paul, S.M. (1997). "Mapping quantitative trait loci for fear-like behaviors in mice." Genomics 46(1):1-8.
46. Gowaty, P.A. et al. (2003). "Male house mice produce fewer offspring with lower viability and poorer performance when mated with females they do not prefer." Animal Behaviour 65: 95-103.
47. Gonzalez-Bernal, E. et al. (2014). "Invasive Cane Toads: social facilitation depends upon an individual's personality." PLOS One 9(7): e102880.

48. Gosling, S. D. (2001). "From mice to men: What can we learn about personality from animal research?" Psychological Bulletin 127(1): 45-86.
49. Gosling, S.D (2008). "Personality in non-human animals." Social and Personality Psychology Compass 2(2):985-1001
50. Gray, S.J. and Hurst, J.L. (1997). "Behavioural mechanisms underlying the spatial dispersion of commensal *Mus domesticus* and grassland *Mus spretus*." Animal Behaviour, 53(3):511-524.
51. Gyuris, E. et al.(2011)."Individual behaviour in firebugs (*Pyrrhocoris apterus*)." Proc R Soc B. 278 (1705): 628-633.
52. Hallatschek, O. and Nelson, D.R. (2008). "Gene surfing in expanding populations." Theor Popul Biol. 73(1):158-170.
53. Hammerschmidt, K. et al. (2009). "Female mice respond to male ultrasonic 'songs' with approach behaviour." Biology Letters 5(5):589-592.
54. Hardouin, E. A. et al. (2015). "Eurasian house mouse (*Mus musculus* L.) differentiation at microsatellite loci identifies the Iranian plateau as a phylogeographic hotspot." BMC Evolutionary Biology 15.
55. Herde, A. and J. A. Eccard (2013). "Consistency in boldness, activity and exploration at different stages of life." BMC Ecology 13.
56. Holmes, A. et al. (2000). "Behavioral profile of wild mice in the elevated plus-maze test for anxiety." Physiology & Behavior 71(5): 509-516.
57. Hurst, J.L. (1987). "The functions of urine marking in a free-living population of house mice *Mus domesticus* Ruddy." Animal Behaviour 35:1433-1442.
58. Hurst, J. L. (1990). "Urine marking in populations of wild house mice *Mus-domesticus* ruddy. 1. Communications between males." Animal Behaviour 40: 209-222.
59. Hurst, J. L. et al. (1998). "Proteins in urine scent marks of male house mice extend the longevity of olfactory signals." Animal Behaviour 55: 1289-1297.
60. Ihle, S. et al. (2006). "An analysis of signatures of selective sweeps in natural populations of the house mouse." Molecular Biology and Evolution 23(4): 790-797.
61. Immelmann, K. (1975). "Ecological significance of imprinting and early learning." Annual Review of Ecology and Systematics 6:15-37.
62. Jang, K.L. et al. (1996). "Heritability of the Big Five personality dimensions and their facets: a twin study." Journal of Personality 64(3):577-592.
63. Jauregi, J. et al. (2013). "Behavioural profile of adults with Prader-Willi syndrome: correlations with environmental variables." J Neurodev. Disord. 5(1):18.
64. Jiang, Y.H. et al. (2010). "Altered ultrasonic vocalization and impaired learning and memory in Angelman syndrome mouse model with a large maternal deletion from *Ube3a* to *Gabrb3*." PLoS One 5:e12278
65. Kalinowski, S. et al. (2007). "Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment." Molecular Ecology 16(5): 1099-1106.
66. Kaufmann, J. H. (1983). "On the definitions and functions of dominance and territoriality." Biological Reviews of the Cambridge Philosophical Society 58(1): 1-20.
67. Kears, M. et al. (2012). "Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data." Bioinformatics 28(12), 1647-1649
68. Klopfer, P.H. (1963). "Behavioural aspects of habitat selection: a preliminary report on stereotypy in foliage preferences of birds." The Wilson Bulletin 77: 376-381.
69. Kluehn, E. and Brommer J.E. (2013). "Context-specific repeatability of personality traits in a wild-bird: a reaction-norm perspective." Behavioural Ecology 24(3):650-658.
70. Knopik, V.S. et al. (2016). "The importance of the prenatal environment in behavioural genetics:introduction to special issue." Behavioural Genetics 46(3):281-285.

71. Koenig, B. et al. (2015). "A system for automatic recording of social behaviour in free-living wild house mouse population." Animal Biotelemetry 3(39): 1-15.
72. Koolhaas, J. M. et al. (1999). "Coping styles in animals: current status in behavior and stress-physiology." Neuroscience and Bio behavioral Reviews 23(7): 925-935.
73. Korsten et al. (2010). "Association between DRD4 gene polymorphism and personality variation in great tits across four wild populations." Molecular Ecology 19(4): 832-843.
74. Kortet et al. (2015). "Personality difference in two minnow populations that differ in their parasitism and predation risk." Front Ecol Evol. 3:9.
75. Krebs, C.J. et al. (1995). "Movements of feral house mice in agricultural landscapes." Australian Journal of Zoology 43:293-302.
76. Kruuk, E.B and Gilchrist S. (1997). "Mechanisms maintain species differentiation: predator-mediated selection in *Bombina* hybrid zone." Proceeding of the Royal Society B 264(1378): 105-110.
77. Labyak, S.E. et al. (1997). "Rhythm chronotypes in a diurnal rodent, *Octodon, degus*." American Journal of Physiology 273(3):R1058-R1066.
78. Lantova, P. et al. (2012). "Urine marking in male common voles: Does behavioural activity matter?" Behavioural Processes 90(2): 174-179.
79. Lenington, S. et al. (1996). "Genetic basis for male aggression and survivorship in wild house mice (*Mus domesticus*)." Aggressive Behavior 22(2): 135-145.
80. Loehlin, J. C. (1982). "Are personality traits differentially heritable?" Behavior Genetics 12(4): 417-428.
81. Le Coeur et al. (2015). "Temporally fluctuating selection on a personality trait in a wild rodent population." Behavioural Ecology 27(6)
82. Lidicker, W.Z. (1975). "Social behavior and density regulation in house mice living in large enclosures." J. Anim. Ecol. 45:677-697.
83. Linnenbrink, M. and von Merten, S. (2017). "No speed dating please! Patterns of social preference in male and female house mice." Frontiers in Zoology 14:38.
84. Lopes, P.C. et al. (2016). "Infection-induced behavioural changes reduce connectivity and the potential for diseases spread in wild mice contact networks" Scientific Reports 6:31790.
85. Lorenc et al. (2014). "Genetic Differentiation of hypothalamus parentally biased transcripts in populations of the house mouse implicate the Prader-Willi Syndrome imprinted region as a possible source of behavioural divergence". Molecular Biology of Evolution 31(12):3240-3249.
86. Lyon, M.F. (2003). "Transmission ratio distortion in mice." Annu. Rev. Genet. 393-408.
87. Malone, N. et al. (2001). "Social status, odour communication and mate choice in wild house mice." Chemical Signals in Invertebrates 9:149-156.
88. Manning, C.J. et al. (1995). "Communal nesting and communal nursing in house mice, *Mus musculus domesticus*." Animal Behaviour 50(3):741-751.
89. Marchetti, C. and P. J. Drent (2000). "Individual differences in the use of social information in foraging by captive great tits." Animal Behaviour 60: 131-140.
90. Martin, J.G.A. and Reale, D. (2008). "Temperament, risk assessment and habituation to novelty in eastern chipmunks." Animal Behaviour 75:309-318.
91. Manser A. et al. (2015). "Female house mice avoid fertilization by t haplotype incompatible males is a mate choice experiment." Journal of Evolutionary Biology 28: 54-64.
92. Meikle D.B and Thornton, M.W. (1995). "Premating and gestational effects of maternal nutrition on secondary sex ratio in house mice." L Reprod Fertil 105(2):159-186.
93. Montero, I. et al. (2013). "Paternal imprinting of mating preferences between natural populations of house mice (*Mus musculus domesticus*)." Molecular Ecology 22(9): 2549-2562.

94. Montero, I. (2010). "Mate choice and reproductive strategies in recently diverged populations of house mouse *Mus musculus domesticus*." Christian Albrechts University zu Kiel, inaugural dissertation.
95. Montiglio, P.O. (2012). "Personality differences are related to long term stress reactivity in a population of wild eastern chipmunks, *Tamias striatus*." Animal Behaviour 84: 1071-1079.
96. Morita T. et al. (1993). "Evolution of the mouse t haplotype." Mechanisms of molecular evolution: Introduction to molecular population biology (Takahata, N. and Clark, A.G. eds) pp. 151-158, Sinauer Associates Inc. Sunderland, MA.
97. Musolf, K. et al. (2010). "Ultrasonic courtship vocalizations in wild house mice, *Mus musculus musculus*." Animal Behaviour 79(3):757-764.
98. Nakagawa, S. and Schielzeth, H. (2010) *Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists*. Biological Reviews 85: 935-956.
99. Newsome, A.E. (1969). "A population study of House-mice temporarily inhabiting a south Australian wheatfield." J. Anim. Ecol. 38: 341-359.
100. Nice, M. M. (1941). "The role of territory in bird life." Amer. Midl. Nat., vol. 26, pp. 441-487.
101. Nicolaus, M. et al. (2016). "Density fluctuations represent a key process maintaining personality variation in a wild passerine bird." Ecology Letters 19(4):478-486.
102. Noyes, R. F. et al. (1982). "Social structure of feral house mouse (*Mus musculus L.*) populations: effects of resource partitioning." Behaviour Ecology and Sociobiology 10:157-163.
103. Oakeshot, J. G. (1974). "Social dominance, aggressiveness and mating success among male house mice (*Mus musculus*)." Oecologia 15(2): 143-158.
104. Patris, B. and Baudoin, C. (2000). "A comparative study of parental care between two rodent species: implications for the mating system of the mound-building mice *Mus spicilegus*" Behav. Proc. 51:35-43.
105. Penke, L. et al. (2007). "The evolutionary genetics of personality." European Journal of Personality 21(5): 549-587.
106. Penn, D.J and Potts, W.K. (1998). "The evolution of mating preferences and major histocompatibility complex genes". The American Naturalist 153(2):145-164.
107. Pinheiro, J. et al. (2016). "*nlme: Linear and Nonlinear Mixed Effects Models*." R package version 3.1-128.
108. Planchart, A. et al. (2000). "Physical mapping male fertility and meiotic drive quantitative trait loci in the mouse t complex using chromosome deficiencies." Genetics 155:803-812.
109. Potts, W. K. et al. (1991). "Mating patterns in semi-natural populations of mice influenced by MHC genotype." Nature 352(6336): 619-621.
110. Pritchard, J. et al. (2000). "Inference of population structure using multilocus genotype data". Genetics 155:945-959.
111. R Development Core Team (2008). "R: A language and environment for statistical computing." R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>
112. Réale, D. and Festa-Bianchet, M. (2003). "Predator-induced natural selection on temperament in bighorn ewes." Anim Behaviour 65:463-470.
113. Reale, D. et al. (2007). "Integrating animal temperament within ecology and evolution." Biological Reviews 82(2): 291-318.
114. Reale, D. (2007). "What do we really know about selection on personality?" European Journal of Personality 21(5): 618-620.
115. Relkovic, D. and Isles, A.R. (2013). "Behavioural and cognitive profiles of mouse models for Prader-Willi syndrome." Brain Res Bull. 92(4):1-8.
116. Revelle, W. (2016). psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, <http://CRAN.R-project.org/package=psych> Version = 1.5.8.

117. Reinhardt, N. (2015). "Territorialverhalten zweier Populationen der Westlichen Hausmaus *Mus musculus domesticus*." Christian Albrechts Universität zu Kiel, Bachelors Thesis.
118. Rimbach, R. et al. (2017). "Contrasting activity budgets of alternative reproductive tactics in male striped mice." Journal of Zoology 301(4):280-289.
119. Roger, S.B. et al. (2013). "Applied spatial data analysis with R", Second edition. Springer, New York. [Http://www.asdar-book-org](http://www.asdar-book-org).
120. Sarkar, Deepayan (2008). "Lattice: Multivariate Data Visualization with R." Springer, New York. ISBN 978-0-387-75968-5
121. Schradin, C. et al. (2010). "Female HRS is regulated by resource distribution and intraspecific competition: a long-term field study." Animal Behaviour 79(1):195-203.
122. Schuett, W. et al. (2009). "Sexual selection and animal personality." Biological Reviews 85:217-246.
123. Schimenti, J. and Hammer, M. (1990). "Rapid identification of mouse t haplotype by PCR polymorphism (PCRP)." Mouse Genome 87: 108.
124. Schoepf, I. et al. (2015). "Manipulation of population density and food availability affects home range sizes of African striped mouse females." Animal Behaviour 99:53-60.
125. Schuster, A.C. et al. (2017). "Repeatability and consistency of individual behaviour in juvenile and adult Eurasian harvest mice." Naturwissenschaften 104(3):10.
126. Selander, R. (1970). "Biochemical polymorphism in population of the house mouse and old-field mouse." Mammal Review 1(2): 39-40.
127. Seppälä, O. and Jokela, J. (2008). "Host manipulation as a parasite transmission strategy when manipulation is exploited by non-host predators." Biology Letters (6): 663-666.
128. Sih, A. et al. (2003). "Behavioural correlations across situations and the evolution of antipredator behaviour in a sunfish-salamander system." Animal Behaviour 65: 29-44.
129. Sih, A. et al. (2004). "Behavioral syndromes: An integrative overview." Quarterly Review of Biology 79(3): 241-277.
130. Skryabin, B. V. et al. (2007). "Deletion of the MBII-85 snoRNA gene cluster in mice results in postnatal growth retardation." PLOS genetics 3(12):e235.
131. Smith, B. R. and D. T. Blumstein (2008). "Fitness consequences of personality: a meta-analysis." Behavioral Ecology 19(2): 448-455.
132. Smith, J M (1982). "Evolution and the theory of games." Cambridge University Press
133. Smiseth, P. T. et al. (2008). "Adaptive consequences and heritable basis of asynchronous hatching in *Nicrophorus vespilloides*." Oikos 117(6): 899-907.
134. Singleton, G.R. (1983). "The social and genetic structure of a natural colony of house mice, *Mus musculus*, at Healesville Wildlife Sanctuary." Australian Journal of Zoology 31:155-166.
135. Singleton, G. R. and D. A. Hay (1983). "The effect of social-organization on reproductive success and gene flow in colonies of wild house mice, *Mus musculus*." Behavioral Ecology and Sociobiology 12(1): 49-56.
136. Stamps, J.A. (2007). "Growth-mortality trade-offs and 'personality traits' in animals." Ecology Letters 10: 355-363.
137. Stamps, J.A. and Groothuis, T.G.G. (2010). "The development of animal personality: relevance, concepts and perspectives." Biological Reviews 85(2):301-325.
138. Taborsky, B. and R. F. Oliveira (2012). "Social competence: an evolutionary approach." Trends in Ecology & Evolution 27(12): 679-688.
139. Tamura, K. et al.(2007) "MEGA8: Molecular Evolutionary Genetics Analysis (MEGA) software version 8.0." Molecular Biology and Evolution 24:1596-1599.
140. Thonhauser, K. et al. (2014). "Multiple paternity does not depend on male genetic diversity". Animal Behaviour 93:135-141.

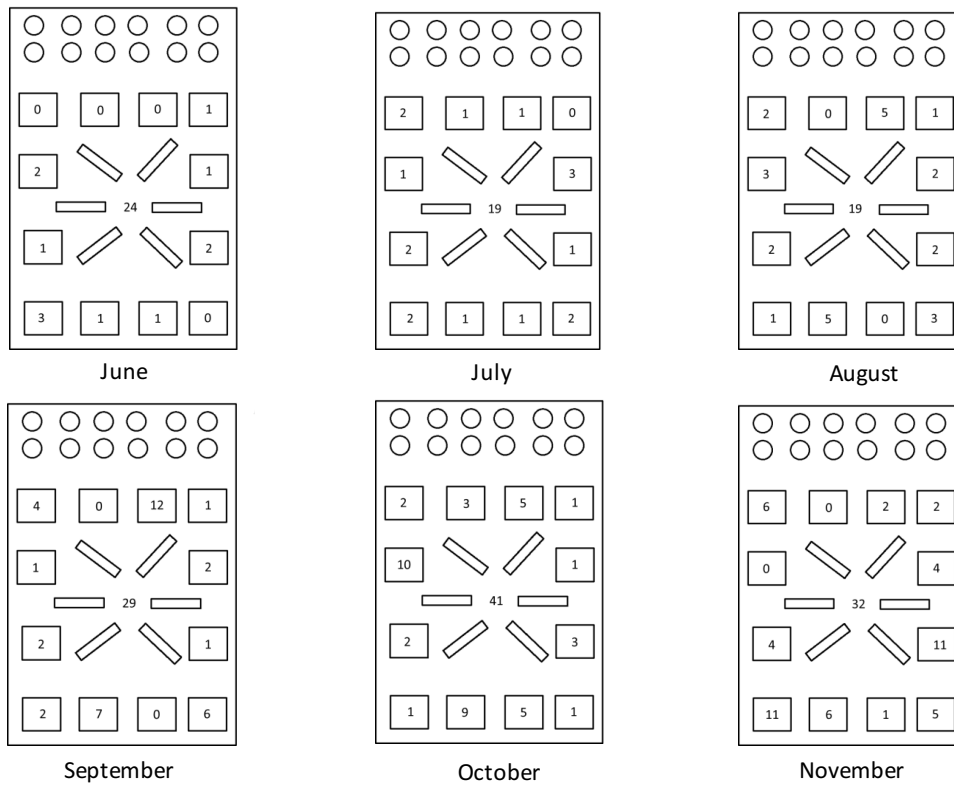
141. Tremmel, M. and Muller, C. (2013). "Insect personality depends on environmental conditions." Behavioural Ecology 24 (2): 386-392.
142. Trivers, R.L. and Willard, D.E. (1973). "Natural selection of parental ability to vary the sex ratio of offspring." Science 179: 90-92.
143. van Oers, K et al. (2004). "Realized heritability and repeatability of risk-taking behaviour in relation to avian personalities." Proceedings of the Royal Society 271(1534): 65-73.
144. van Oers, K. et al. (2008). "Personality is associated with extrapair paternity in great tits, *Parus major*." Animal Behaviour 76: 555-563.
145. van Oers, K. and Mueller, J. (2011). "Evolutionary genomics of animal personality." Philos Trans R Soc Lond B Biol Sci 365(1560): 3991-4000.
146. van Oortmerssen, G. A. and T. C. M. Bakker (1981). "Artificial selection for short and long attack latencies in wild *Mus musculus domesticus*." Behavior Genetics 11(2): 115-126.
147. Veenema, A. H. et al. (2003). "Genetic selection for coping style predicts stressor susceptibility." Journal of Neuroendocrinology 15(3): 256-267.
148. Verbeek, M. E. M. et al. (1994). "Consistent individual differences in early exploratory-behaviour of male great tits." Animal Behaviour 48(5): 1113-1121.
149. Verbeek, M. E. M. et al. (1999). "Individual behavioural characteristics and dominance in aviary groups of great tits." Behaviour 136: 23-48.
150. von Merten, S. et al. (2014). "A Role for Ultrasonic Vocalisation in Social Communication Divergence of Natural Populations of the House Mouse (*Mus musculus domesticus*)." Plos One 9(5).
151. Wecker, S. (1963). "The role of early experience in habitat selection by the prairie deer mouse, *Peromyscus maniculatus bairdi*." Ph.D Thesis, University of Michigan.
152. Weidt, A. et al. (2014). "Communal nursing in wild house mice is not a by-product of group living: Females choose." Naturwissenschaften 101(1):73-76.
153. Wickham, H. (2009). ggplot2: elegant graphics for data analysis. Springer New York
154. Wilson, R. C. et al. (1976). "Open-field behaviour in Muroid rodents." Biology 17(4): 495-506. Wilson DS. (1998) "Adaptive individual differences within single populations." Philos Trans R Soc Lond B Biol Sci 353:199-205.
155. Wilson, D.S. et al. (1994) "Shyness and boldness in humans and other animals." Trends Ecol Evol 19:442-446.
156. Wright, S.L et al. (1988). "Allocation of reproductive effort in *Mus domesticus*: responses of offspring sex ratio and quality to social density and food availability." Behavioural Ecology and Sociobiology 23(6):357-365.
157. Wolak, D.J. et al. (2012) Guidelines for Estimating Repeatability. Methods in Ecology and Evolution 3(1):129-137.
158. Wolff, R.J. (1985). "Mating behaviour and female choice: their relationship to social structure in wild caught house mice (*Mus musculus*) housed in a semi-natural environment." Journal of Zoology (London) 207:43-51.
159. Wolff, J. O. and P. W. Sherman (2007). Rodent societies as model systems, University of Chicago Press.
160. Wolff, R.O. (2009). "Mating behaviour and female choice: Their relation to social structure in wild caught House mice (*Mus musculus*)." Journal of Zoology 207: 43-51.
161. Young, R. and D. N. Johnson (1991). "A fully automated light dark apparatus useful for comparing anxiolytic agents." Pharmacology Biochemistry and Behavior 40(4): 739-743.
162. Yuen, C. H. et al. (2015). "Personality does not constrain social and behavioural flexibility in African striped mice." Behavioral Ecology and Sociobiology 69(8): 1237-1249.

7. Supplemental Information

Comments about technical issues in enclosures:

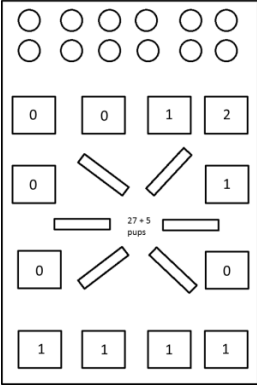
For future experiments, I believe it is important to keep the following things in mind: first; that the antennae on the water bottles are often used as “shelters” by the mice. When the water bottles were changed, the water bottles that were emptied the fastest were the ones at the centre of the enclosure and with the apparent fewest reads. This suggests the other bottles might have been guarded, even if it was unintentional, by the mice sleeping under them. Second; some mice chose to live in the cables or outside the nest boxes. There seemed to be one such family in every enclosure. In some cases, these mice would even pull bedding out of nest boxes, to build nests by the door (in the area with the food and water). These families also bred well, and when we entered the enclosure, they simply climbed into the cables or dispersed while we worked in the enclosure, and returned to the “nest” when we closed the door. In one enclosure, the mice made a hole in the wall and started to nest there. Before the cable guards were added, mice also tried to nest above the door. Finally, there was a delay in the production of offspring after the mice were put in the experiment. When we tried to increase the length of the experiment to ensure more first and second generation mice were born, we got many types of backcrosses. This appears to happen only at the beginning of the experiment and extending for this purpose alone is not helpful. Finally, now that we have established that the measurements are repeatable, it would be possible to use these tests to phenotype mice at the beginning of the experiment in future tests, to minimize disruption to the enclosures in future projects.

Enclosure A

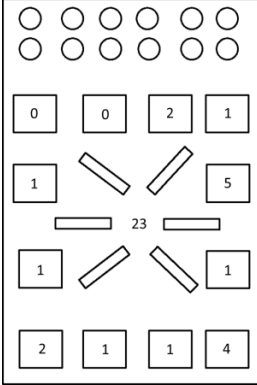


Supplemental Figure S.1.A: Monitoring events in enclosure A. The number of mice caught free and in houses from June until November in enclosure A.

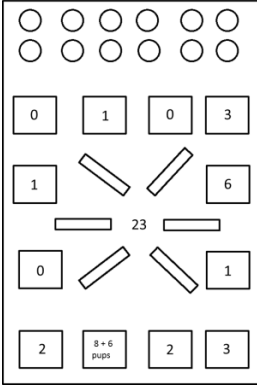
Enclosure B



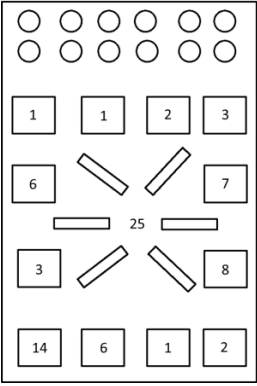
June



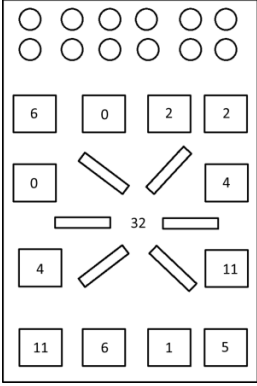
July



August

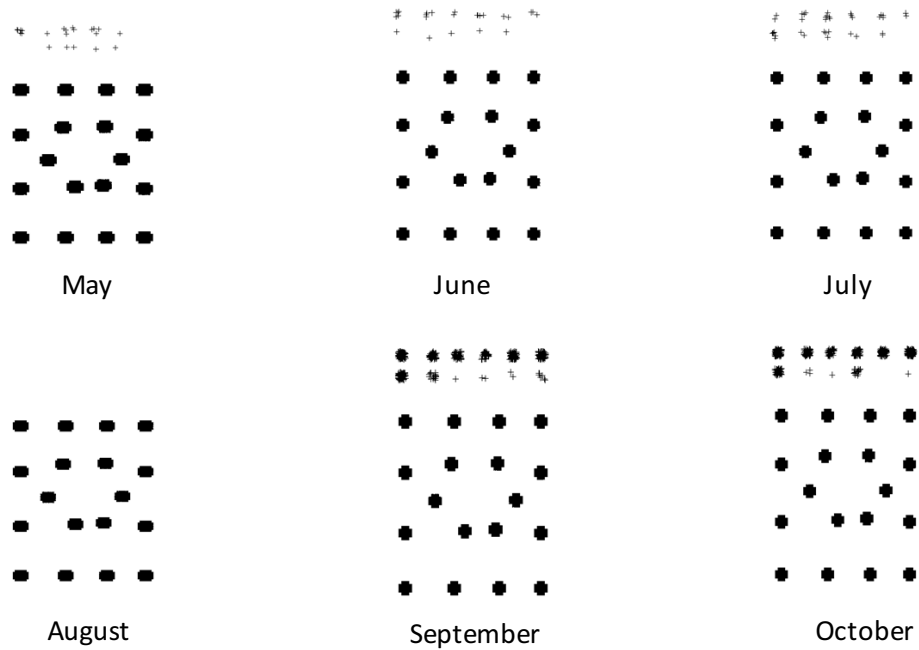


September

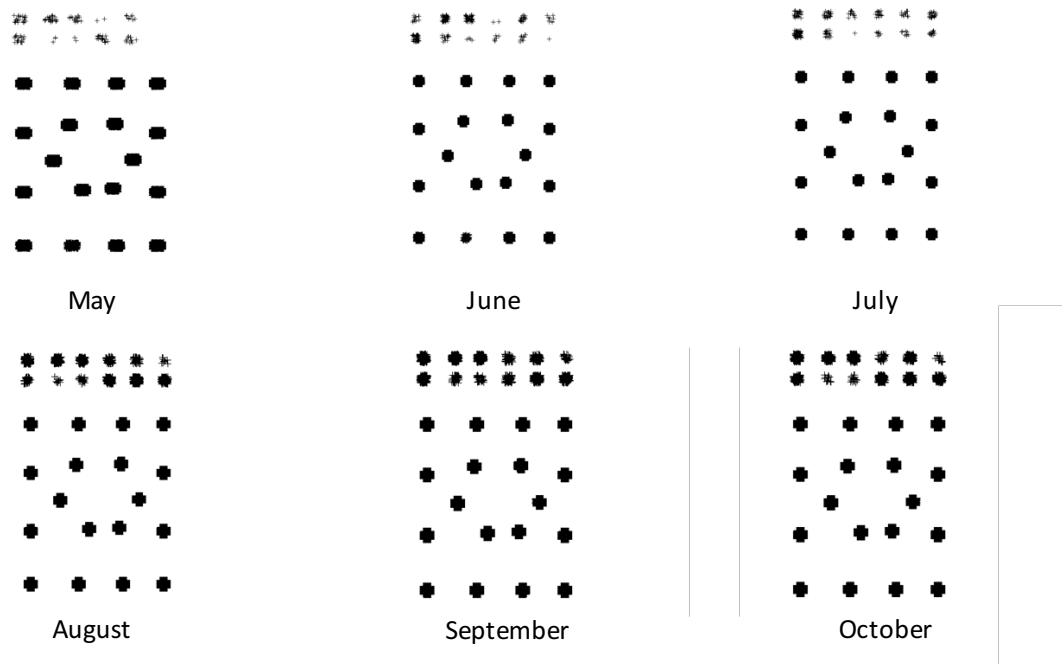


October

Supplemental Figure S.1.B: Monitoring events in enclosure B. The number of mice caught free and in houses from June until October in enclosure B.



Supplemental Figure S.2.A: Antenna reads enclosure A. Antenna reads by month in enclosure A. In enclosure A in August there was a power failure in the first loop (water bottles) and no antenna reads from this area.



Supplemental Figure S.2.B: Antenna reads enclosure B. Antenna reads by month in enclosure B.

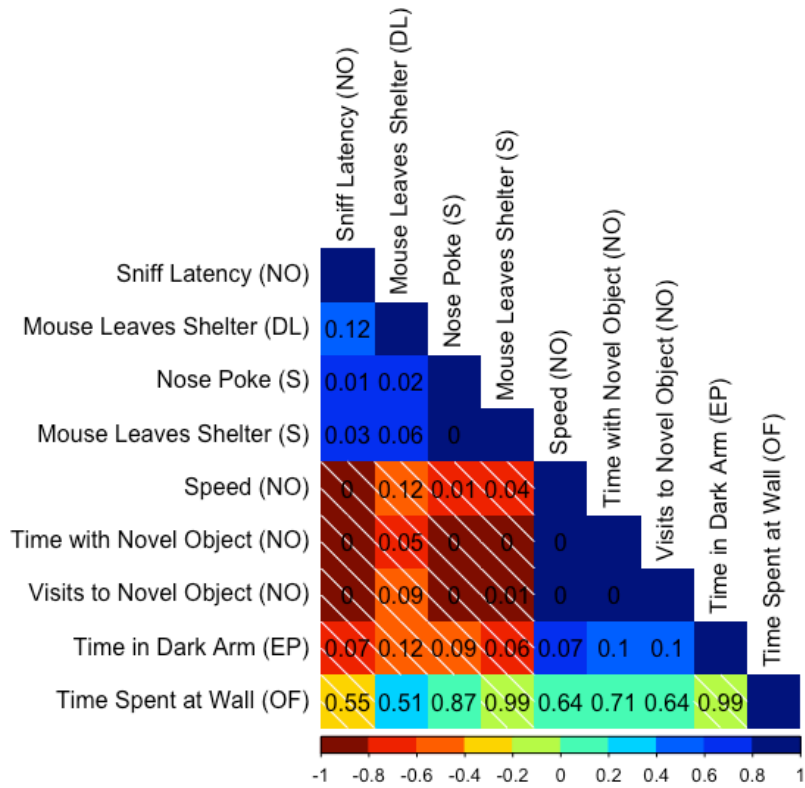


Figure S.3: Correlation matrix with Holm Correction. Correlation plot for nine measurements from four different behavioural tests which were repeatable across the course of the experiment. Novel Object (NO), Dark/Light box (DL), Startle Test (S), Novel Object Test (NO), Elevated Plus Maze (EP) and Open Field Test (OF) (n=33).

Digital supplement

Experimental excel sheets and individual territory maps will be made available in a digital archive. A description of the tables and the territory maps for individuals are described below.

Overview of Tables:

Loci	All samples typed by sample
F0_Behaviour_All	F0 behaviours and spatial use scores
Activity_Enclosures	Activity scores from enclosures
MonitoringRoom5T2	Results of 8 monitorings in enclosure A
MonitoringRoom13T2	Results of 8 monitorings in enclosure B
MonitoringRoom14T2	Results of 8 monitorings in enclosure C
Observed Props	1. Observed G1xG1 mating proportions 2. Mating List 3. Expected G1xG1 mating based on mice present at the time
Parent to Offspring Territory Scores	Offspring and Parents territory scores

Folders MCPs Enclosures A, B and C show territory use maps for the entire course of the experiment.

Declaration/Erklärung

I hereby declare that:

- i. Apart from my supervisor's guidance, the content and design of this dissertation is the product of my own work. The co-authors contributions to specific paragraphs are listed in the thesis outline section;
- ii. This thesis has not already been submitted either partially or wholly as part of doctoral degree to another examination body, and no other materials are published or submitted for publication than indicated in the thesis;
- iii. The preparation of the thesis has been subjected to the Rules of Good Scientific Practice of the CB/G Research Foundation.

Rebecca Krebs-Wheaton

Plön, October 8th, 2017

Name Rebecca Krebs, geb. Wheaton

Anschrift Achterkamp 106
24106 Kiel

Geburtsdaten 19.06.1986 in Moncton, NB, Kanada

Staatsangehörigkeit Kanadisch

Hochschulausbildung

2004-2008
Bachelor of Science
University of Prince Edward Island,
Charlottetown, PEI, Kanada

2010-2013
Master of Science in Ecology and Evolution
Universität Bern, Schweiz
Abteilung: Populationsgenetik
Betreuung durch: Dr. Gerald Heckel

2013-jetzt
Promotion in Zoologie
Max-Planck-Institut für Evolutionsbiologie,
Plön, Deutschland
Abteilung: Evolutionsgenetik
Betreuung durch: Diethard Tautz und
Miriam Linnenbrink

Date of Oral Examination
Voraussichtlicher Abschluss der Promotion
Mathematisch-Naturwissenschaftliche
Fakultät Christian-Albrechts-Universität zu
Kiel