

# Conclusions from a behavioral aging study on male and female F2 hybrid mice on age-related behavior, buoyancy in water-based tests, and an ethical method to assess lifespan

Julia Adelöf<sup>1,2</sup>, Jaime M. Ross<sup>3,4</sup>, Stanley E. Lazic<sup>5,6</sup>, Madeleine Zetterberg<sup>1</sup>, John Wiseman<sup>2</sup>, Malin Hernebring<sup>1,2</sup>

<sup>1</sup>Department of Clinical Neuroscience, Institute of Neuroscience and Physiology, Sahlgrenska Academy at the University of Gothenburg, Gothenburg 41390, Sweden

<sup>2</sup>Discovery Biology, Discovery Sciences, R&D AstraZeneca, Gothenburg, Mölndal 43153, Sweden

<sup>3</sup>Department of Genetics, Paul F. Glenn Center for the Biology of Aging, Harvard Medical School, Boston, MA 02215, USA

<sup>4</sup>Department of Neuroscience, Biomedicum, Karolinska Institutet, Stockholm 17165, Sweden

<sup>5</sup>Quantitative Biology, Discovery Sciences, R&D AstraZeneca, Cambridge CB4 0WG, UK

<sup>6</sup>Current address: Prioris.ai Inc., Ottawa K2P 2N2, Canada

**Correspondence to:** Malin Hernebring; email: [malin.hernebring@gu.se](mailto:malin.hernebring@gu.se)

**Keywords:** F2 hybrid mice, aging, sex comparison, exploratory activity, water-based behavioral tests

**Abbreviations:** DEXA: dual energy X-ray absorptiometry; BMC: bone mineral content; BMD: bone mineral density; PAT: passive avoidance test

**Received:** May 23, 2019

**Accepted:** August 22, 2019

**Published:** September 11, 2019

**Copyright:** Adelöf et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

## ABSTRACT

Due to strain-specific behavioral idiosyncrasies, inbred mouse strains are suboptimal research models for behavioral aging studies. The aim of this study is to determine age-related behavioral changes of F2 hybrid C57BL/6NxBALB/c male and female mice. Lifespan was followed ( $n_{\text{males}}=48$ ,  $n_{\text{females}}=51$ ) and cohorts of mature adult (7 months), middle-aged (15 months), and old mice (22 months of age;  $n=7-12$  per group) were assessed regarding open-field activity, exploration, passive avoidance learning/memory, and depressive-like behavior. We found that both males and females demonstrated decreased exploratory behavior with age, while memory and depressive-like behavior were maintained. Females exhibited enhanced depressive-like behavior compared to males; however, a correlation between fat mass and swimming activity in the test directly accounted for 30-46% of this behavioral sex difference. In addition, we suggest a method to qualitatively estimate natural lifespan from survival analyses in which animals with signs of pain or severe disease are euthanized. This is, to our knowledge, the first behavioral study to consider both sex and aging in hybrid mice. We here define decreased exploratory behavior as a conserved hallmark of aging independent of sex, highlight the effect of buoyancy in water tests, and provide a method to assay lifespan with reduced animal suffering.

## INTRODUCTION

Mice are the leading mammalian model system for studying genetic effects on cognitive function and are well-suited model organisms for gerontological research with their relatively short lifespan and economic

husbandry. Traditionally, mouse studies are conducted with inbred strains with the advantage of having less variation by genetic homogeneity and stable characteristics [1]. However, inbred strains can be problematic for aging research since they develop strain-specific maladies with advancing age and thus

studies that intend to determine anti-aging effects might only pick up processes specifically targeting these conditions. Inbred strains also demonstrate age-related behavioral idiosyncrasies, *e.g.* coordination, learning capacity, and anxiety-like behavior, which can limit findings to the specific strain used [2, 3]. To ensure that these factors do not affect the results, it is considered favorable to use hybrid mice [4–6]. F2 hybrids are genetically similar but never uniformly homozygous, which reflects heterogenetic populations better than inbred mice and improves the extrapolation from mice to humans [5, 7].

There is a male sex bias in all biomedical disciplines [8, 9], although several studies confirm a sex difference in behavioral testing of mice [10–12]. Male rodents are dominant models of pharmaceutical discovery and testing, including several anxiolytic and anti-depressive drugs, despite the majority of recipients being women [13, 14]. Additionally, there are known sex differences in uptake and effect of psychotropic drugs, as well as symptoms and adverse side effects in humans [15–17]. Female exclusion has been rationalized by menstrual fluctuations interfering with behavioral data; however, a number of studies have shown that while the estrous cycle increases variability within female cohorts, behavioral differences between the sexes are independent of estrous cycle effects [18, 19].

Behavioral studies consisting of several complimentary tests allow for stronger phenotypic interpretations [20–22]. General behavior, locomotor activity, and exploratory behavior are easily assessed by open-field testing. Activity in the open-field and especially exploratory behavior of both male and female C57BL/6J mice [23–26] has been shown to decline with age and can fundamentally influence other behavioral testing during the aging process [23].

A substantial body of literature reports an aging effect in learning and memory, predominately assessed by spatial tests like Morris water, radial, and Barnes mazes [as reviewed in 27–29]. Decreased performance in these tests has been reported to correlate with female estrous cycle decay and aging [10, 26, 30, 31]. Additionally, sex differences in learning and memory have been found with spatial, cued, and water-based cognition tests, in which females performed poorer than males [30, 32–35]. Spatial reference tests, however, are unsuitable for sex comparisons in rodents since females and males are known to use different cues for navigation, and thus different parameters of these spatial tasks often favor either sex [36–39]. Interestingly, females and males perform equally well in non-spatial learning and memory tests, such as object memory consolidation [23, 30] and active avoidance tests [6, 10, 35].

The forced swim test is one of the most commonly used tests to assess depressive-like behavior by recording the activity of mice placed in water tanks. Immobility is considered a measurement of despair and is frequently used for anti-depressant screening in mice [40–43]. In addition to age [26], several factors such as strain, sex, and handling of mice can influence forced swim test performance [44, 45].

In this work, we analyzed the behavior of male and female C57BL/6N×BALB/c F2 hybrid mice as they age, by following the lifespan of littermates ( $n_{\text{males}}=48$ ,  $n_{\text{females}}=51$ ) and subjecting cohorts of mature adult (7 months), middle-aged (15 months), and old mice (22 months of age;  $n=7-12$  per group) to behavioral phenotyping. To our knowledge, this is the first behavioral study to consider both sex and aging in mice with a hybrid background.

## RESULTS

### An estimation of C57BL/6N×BALB/c F2 natural lifespan

C57BL/6N×BALB/c F2 female and male hybrid mice were followed in a lifespan study, strictly following Swedish animal ethics regulations; if an animal displayed signs of pain or severe disease the animal was euthanized. The fate of the animals in the study is displayed in Table 1. The three shortest-lived and longest-lived animals are included as an indication of onset of death and maximum lifespan. As depicted, 15% of the males and 25% of the females were found dead in the cage without previous signs of pain or severe disease and thus died from intrinsic causes, generally defined as natural deaths in survival analyses. No autopsies were made but brief descriptions of extraordinary physical features were noted (*e.g.* enlarged spleen and various tumors; Supplementary Table 1).

For lifespan analysis, two survival curves per sex were made. In Survival Curve A, euthanization was counted as natural death, while in Survival Curve B euthanized animals were considered healthy when removed from the study. Since animals of severe disease are likely to live longer than the onset of disease, but not as long as healthy littermates, Survival Curve A is an underestimation of natural lifespan and Survival Curve B is an overestimation. Taking both curves into account provides an interval of natural lifespan (Figure 1). Data from Survival Curves A and B are presented in Table 2. There were no statistically significant differences when comparing lifespan of male and female F2 hybrids, though trends of female longevity over male were

**Table 1. C57BL/6N×BALB/c F2 lifespan experiment.**

	Number of animals	Part of cohorts	Euthanized	Natural deaths	% Natural death	Onset of death*	3 most long lived
♂ Survival	48	28	13	7	15	311/312/613	860**/911**/919
♀ Survival	51	29	9	13	25	352/406/560	966/1001/1051

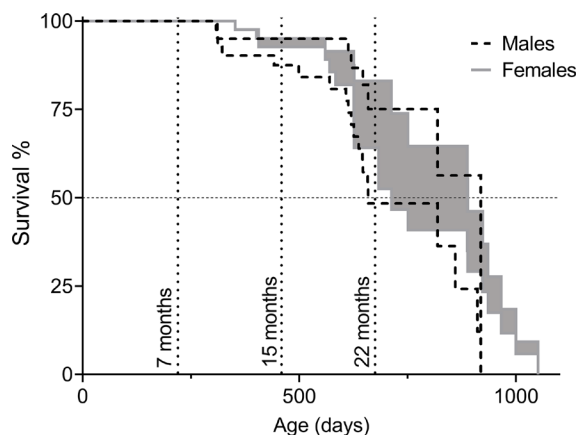
\*3 first animals to die from natural causes

\*\*Removed due to pain or severe disease

observed on mean and median lifespan, as well as on the age of the three longest-lived animals (Table 1).

### Physiological analyses demonstrated that females have less lean mass and more body fat than males

Cohorts of mice were analyzed for behavioral phenotyping at 7, 15, and 22 months of age representing mature adults, middle-aged, and old. The mice were naïve before behavioral testing at all timepoints. As shown in Table 2, no survival loss had yet occurred at the 7-month timepoint and the mice were considered healthy adults. At the 15-month timepoint, survival was around 90% and for the final behavioral analyses at 22 months, 71% of the male and 74% of the female mice remained.



**Figure 1. Survival of C57BL/6N×BALB/c F2 hybrid male and female mice presented as intervals of natural lifespan.** The lower limit of the lifespans, referred to as “Survival Curve A” in Table 2, was obtained by considering euthanization of animals as the same fate as natural death and the upper limit of the lifespans, referred to as “Survival Curve B” in Table 2, by considering euthanized animals as healthy upon removal (censored) and only animals that died of intrinsic causes were counted as deaths. Animals in the cohorts for behavioral assessment were included in the survival analysis until the time of their first test (7, 15, or 22 months of age). Timepoints are indicated by dotted vertical lines. Total number of animals:  $n_{\text{males}}=48$ ,  $n_{\text{females}}=51$ .

Physiological parameters of the mice in the cohorts were assessed to validate for general health status and to identify factors that may influence behavioral analyses. As shown in Figure 2A, average body weight did not differ significantly between 7, 15, and 22 months of age or between sexes. Despite similar body weights, there were clear changes in body composition. DEXA analyses demonstrated that males had higher percent lean mass (Figure 2B;  $p_7=0.0078$ ,  $p_{15}=0.0011$ ) and less percent fat mass (Figure 2C;  $p_7=0.0023$ ,  $p_{15}=0.0012$ ), as compared to females. Percent lean mass also increased with age (Figure 2B;  $p_{M7-22}=0.040$ ,  $p_{M15-22}=0.036$ ,  $p_{F15-22}=0.0098$ ). At 15 months of age, the most distinct sex differences observed were an average of 11.8% more fat mass and 10.5% less lean mass in females as compared to males. Females also continuously displayed a higher core body temperature, as compared to males (Supplementary Table 2;  $p_7<0.0001$ ,  $p_{15}=0.0044$ ,  $p_{22}=0.00013$ ). Neither bone mineral content (BMC) nor density (BMD) declined for either sex at the last timepoint (Supplementary Table 2).

### Exploratory behavior, but not necessarily activity, decreased with age

The activity box, an open-field test, was used to analyze horizontal and rearing activities of naïve hybrid female and male mice. A mouse standing on its hind limbs, termed rearing, is considered both exploratory and general vertical locomotor activity [46]. Data were separated into the first 5 minutes, representing the initial exploratory phase of the test, and the full one-hour to enable analysis of general activity and behavior in a more acquainted environment, also known as intrasessional (within session) activity. An age-effect was detected during the exploratory phase of the open-field session, with horizontal and rearing activity clearly declining from 7 to 22 months in both sexes (Figure 3A; horizontal  $p_{M7-22}=0.0065$ ,  $p_{F7-22}=0.0006$ ; rearing  $p_{M7-22}=0.0003$ ,  $p_{F7-22}=0.0020$ ). The one-hour test results demonstrated that horizontal activity did not change comparing the 7-month and the 22-month age groups of both males and females (Figure 3B; the activity was maintained through 15 months of age in males, while females exhibit an activity drop that may be caused by

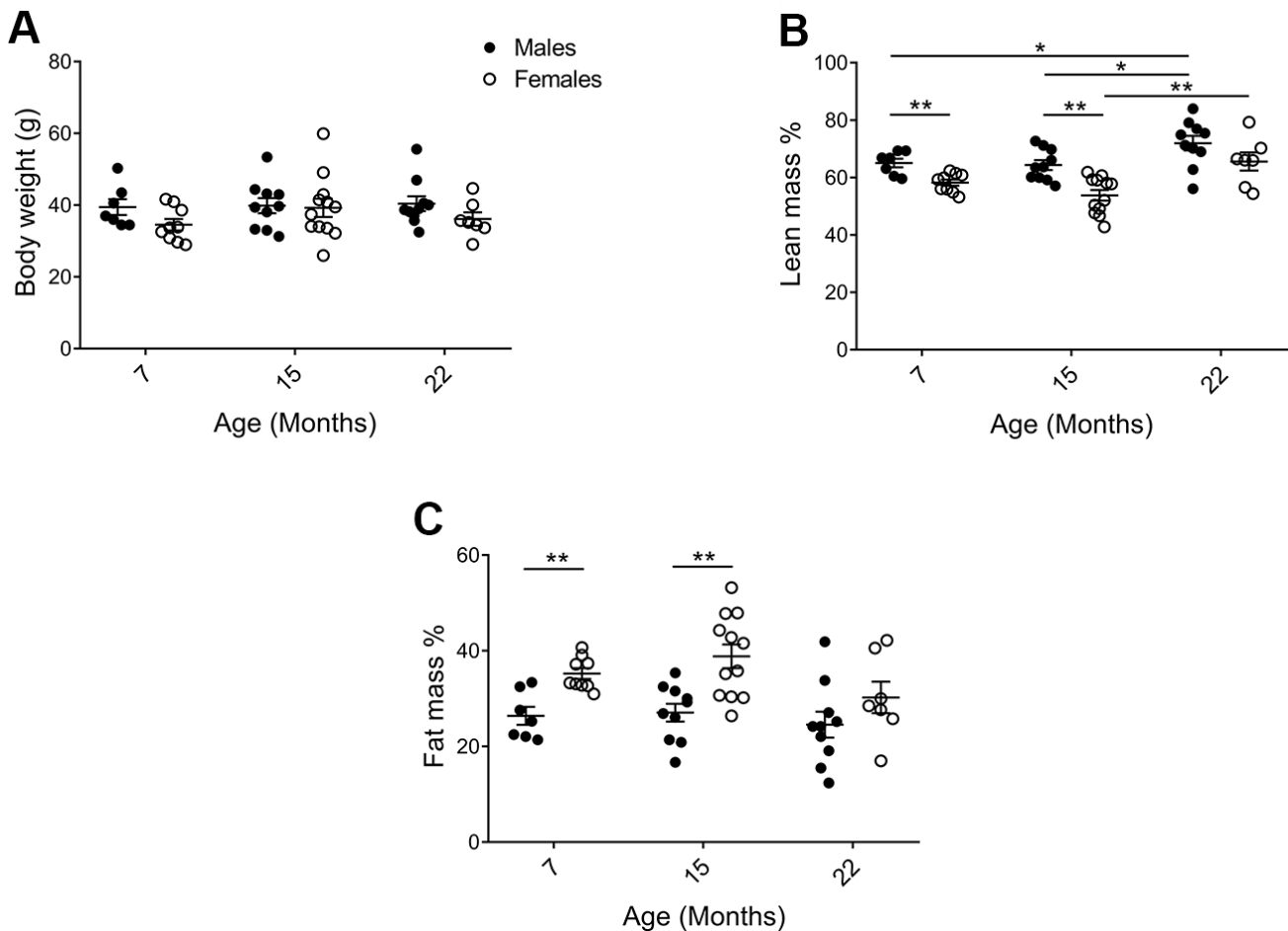
**Table 2. Lifespan analysis C57BL/6N×BALB/c F2 to generate interval of natural lifespan in days.**

	Mean lifespan*	50% survival	7 months' survival**	15 months' survival**	22 months' survival**
♂ Survival curve A	708 ±34	659	100%	85%	60%
♂ Survival curve B	812 ±34	919	100%	95%	83%
♂ <b>Mean A&amp;B</b>	<b>760 ±52</b>	<b>789 ±130*</b>	<b>100%</b>	<b>90 ±5%*</b>	<b>71 ±11%*</b>
♀ Survival curve A	754 ±36	713	100%	92%	64%
♀ Survival curve B	851 ±38	889	100%	95%	83%
♀ <b>Mean A&amp;B</b>	<b>803 ±48</b>	<b>801 ±88*</b>	<b>100%</b>	<b>93 ±2%*</b>	<b>74 ±10%*</b>

In Survival Curve A, euthanized animals were counted as if they had died from natural causes at the time of euthanization, and in Survival Curve B, euthanized animals were considered healthy at that timepoint. This gives the interval of natural lifespan “Mean A&B”.

\*Values are mean ± SEM for Survival Curve A and Survival Curve B, and for Mean A&B variation is depicted as the span between Survival Curve A and B.

\*\*Extrapolated from curve fit

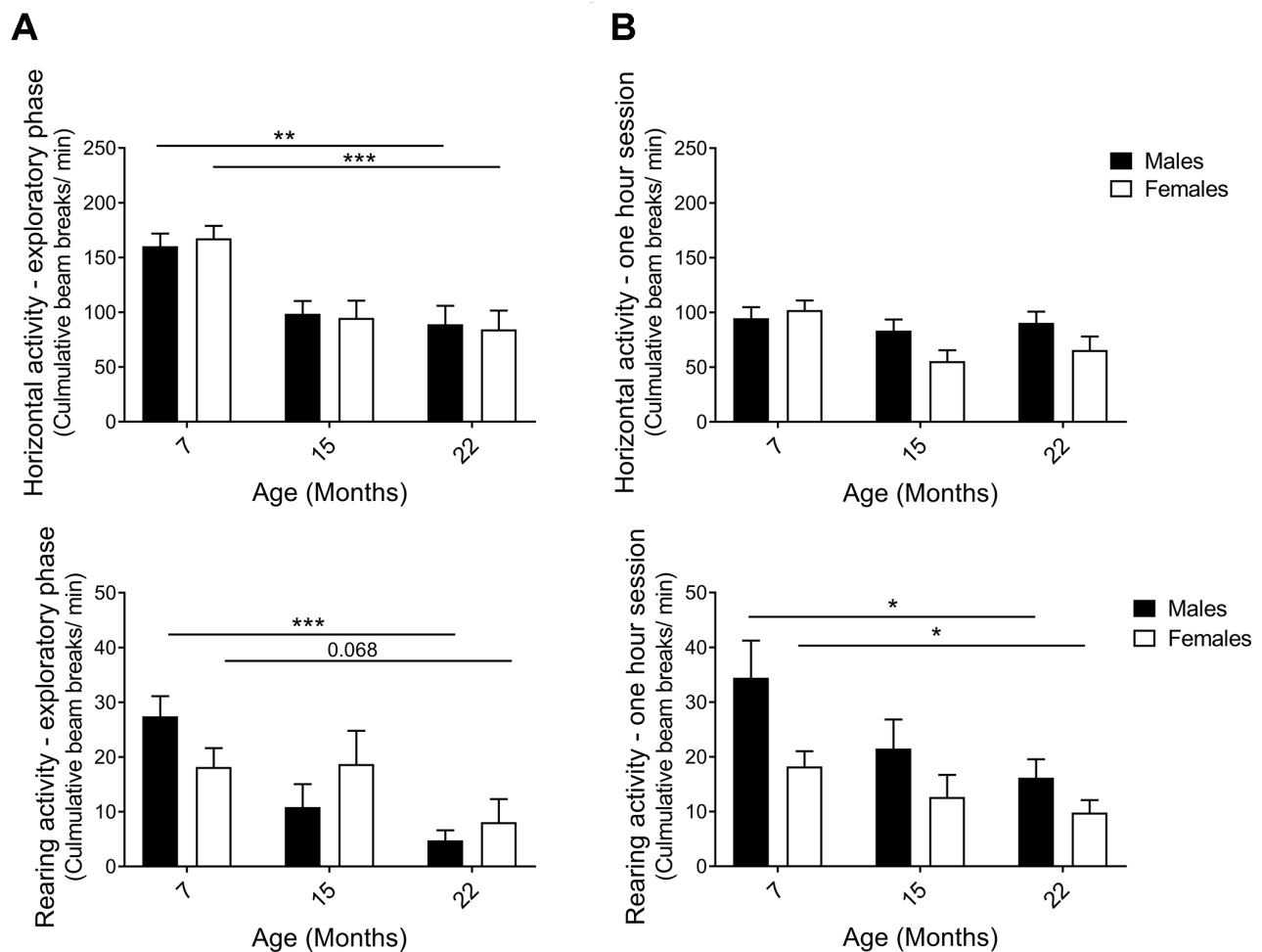


**Figure 2. Physiological parameters of mice in the cohorts for behavioral assessment.** (A) There was no difference in body weight between the cohorts, but (B) DEXA analysis showed that males had higher percentage lean mass than females ( $p_{M7-F7}=0.0078$ ,  $p_{M15-F15}=0.0011$ ; and lean mass increased with age in males  $p_{M7-22}=0.040$ ,  $p_{M15-22}=0.036$ ; Mann Whitney), whereas (C) females exhibited increased fat mass as compared to males ( $p_{M7-F7}=0.0023$ ,  $p_{M15-F15}=0.0012$ ; Welch’s t-test). Values are mean ± SEM;  $n_M=7$ ,  $n_{M15}=10$ ,  $n_{M22}=10$ ,  $n_{F7}=9$ ,  $n_{F15}=12$ ,  $n_{F22}=7$ .

different handling of this particular group due to technical reasons, see Methods). Moreover, horizontal activity was found to decline within session as time elapsed in the open-field and the novelty of the environment decreased (Supplementary Figure 1A). Rearing activity over the full hour also declined with age in both males and females (Figure 3B;  $p_{F7-22}=0.043$ ,  $p_{M7-22}=0.043$ ). In contrast to horizontal activity, no differences in rearing within session were seen, *i.e.* no intrasessional changes (Supplementary Figure 1B). Thus, activity in the explorative phase was more significant and uniform in exhibiting an age-related decline than the activity during the longer sessions. Taken together, diverse exploratory behaviors decreased consistently with age in both sexes.

### Hybrid mice exhibited unaltered learning and memory with aging in the shuttle box passive avoidance test

The shuttle box passive avoidance test is a two-day procedure to assay hippocampal and amygdala-dependent learning and memory through pain conditioning [25, 30, 47, 48]. On training day, mice were allowed to enter a dark compartment, which subsequently resulted in a small electric shock. On the following day, the difference in time to re-enter the dark compartment compared to the previous day was recorded as a measure of recollection of the unpleasant experience. In order to clearly illustrate the memory response, results are presented as time to enter on Day 2



**Figure 3. Horizontal and rearing activity per minute in an activity box open field test. (A)** Activity in the first 5 minutes – the exploratory phase – decreased with aging for both males and females. The decline in horizontal ( $p_{M7-22}=0.0065$ ,  $p_{F7-22}=0.0006$ ; Student's t-test) and rearing activities ( $p_{M7-22}=0.0003$ ,  $p_{F7-22}=0.068$ ; Mann-Whitney) indicated that exploratory behavior was highly affected by age. **(B)** Activity in 1-hour sessions of the same test. Horizontal activity was not altered from 7 to 22 months of age, whilst rearing activity declined for both female and male hybrids, thus illustrating an aging effect ( $p_{F7-22}=0.043$ ,  $p_{M7-22}=0.043$ ; Mann-Whitney). The decrease in horizontal activity observed in 15-month-old females was likely caused by different handling of this particular group due to technical reasons (see Methods). Values are mean  $\pm$  SEM;  $n_{M7}=7$ ,  $n_{M15}=10$ ,  $n_{M22}=10$ ,  $n_{F7}=10$ ,  $n_{F15}=12$ ,  $n_{F22}=7$ .

subtracted by time to enter on Day 1. Although 7-month-old males tended to keep away from the avoidance-trained area the longest, indicating the best learning and memory of all groups, large individual variations were observed within the cohorts and no statistical significances were found. Thus, we unexpectedly found no trends of declining learning and memory with age in neither sex (Figure 4).

### Increased immobility of females in forced swim test which is partially due to higher fat mass

Immobility in the forced swim test is an indicator of depressive-like behavior. We did not find any age-related differences in forced swim test performance in males or females (Figure 5A;  $p=0.244$ ; 2-way ANOVA); however, when comparing males and females we observed a clear effect of sex on immobility ( $p=0.001$ ; 2-way ANOVA). Since there was little evidence for a sex by age interaction ( $p=0.267$ ; 2-way ANOVA), the three age groups were combined in subsequent analyses. Comparing pooled age groups demonstrated that females had around 21% increased immobility compared to males (Figure 5B;  $p=0.0002$ ; Mann-Whitney).

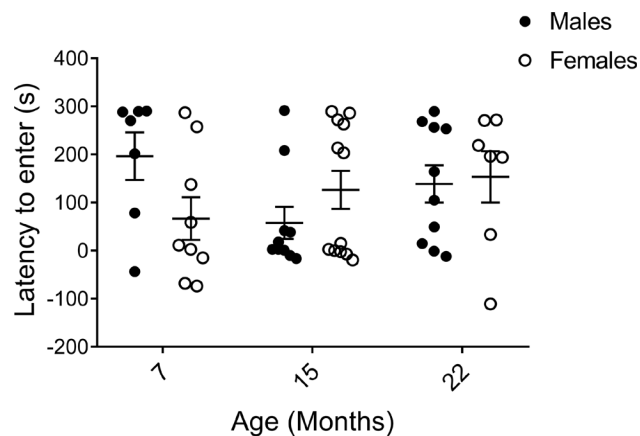
We controlled whether body weight affected performance in all our behavioral tests but could not find any correlations. However, with females having a greater proportion of body weight comprised of fat (Figure 2C) they are likely to have increased buoyancy which may affect their performance in the forced swim

test. Indeed, for both males and females, animals with a higher percent fat mass were less active (Figure 5C). We therefore used a Bayesian mediation analysis to test if, and to what extent, sex differences in the forced swim test were mediated by fat mass, and/or via other mechanisms [49, 50]. We found that fat mass accounted for about 30% of the sex effect (Figure 5D; Bayesian  $p=0.96$ ; a Bayesian  $p$ -value close to 1.0 indicates that the conclusion is likely), while the remaining 70% was caused by other mechanisms (Figure 5D; Bayesian  $p=0.99$ ). The differences in fat mass and immobility between males and females were most distinct at 7 and 15 months of age, despite no significant age, or age to sex interaction effect. As forced swim tests are generally not performed on aged animals, we performed the same Bayesian mediation analysis on the data set excluding the 22-month-old cohort, and found that 46% of the sex effect is caused by fat mass in this data set (Supplementary Figure 2; Bayesian  $p=0.986$ ), while the difference in immobility between males and females was in the same range (24% compared to 21%). Hence, we can conclude that fat mass accounted for 30%-46% of the sex effect in the forced swim test, depending on the age groups examined.

## DISCUSSION

In this study, F2 hybrid female and male mice were assessed for behavioral tests with the aim to investigate sex differences and age-related alterations. Increasing our understanding of age-related changes in mouse behavior is vital for successful research on healthy aging. We herein: 1) reveal decreased exploratory behavior, but not altered learning/memory or depressive-like behavior, as a robust behavioral marker of aging in both male and female hybrid mice, 2) did not find sex differences in learning and memory assessed with a passive avoidance test, 3) couple immobility in a water test to body composition and demonstrate that fat mass accounted for 30-46% of the observed increase in depressive-like behavior of females compared to males, and 4) present a novel method to estimate the natural lifespan from survival studies in which animals in pain or with severe disease are euthanized.

Traditionally, lifespan studies are conducted until all animals die from natural causes and euthanization is only considered when end-stage diseases make survival for more than an additional week highly unlikely. Since aging is associated with many severe diseases, these types of experiments are especially challenging to combine with ethical considerations of animal welfare. Still, to be able to compare studies on aging phenotypes, it is crucial to distinguish the lifespan timepoint of the cohort analyzed. Here, we present a novel method that

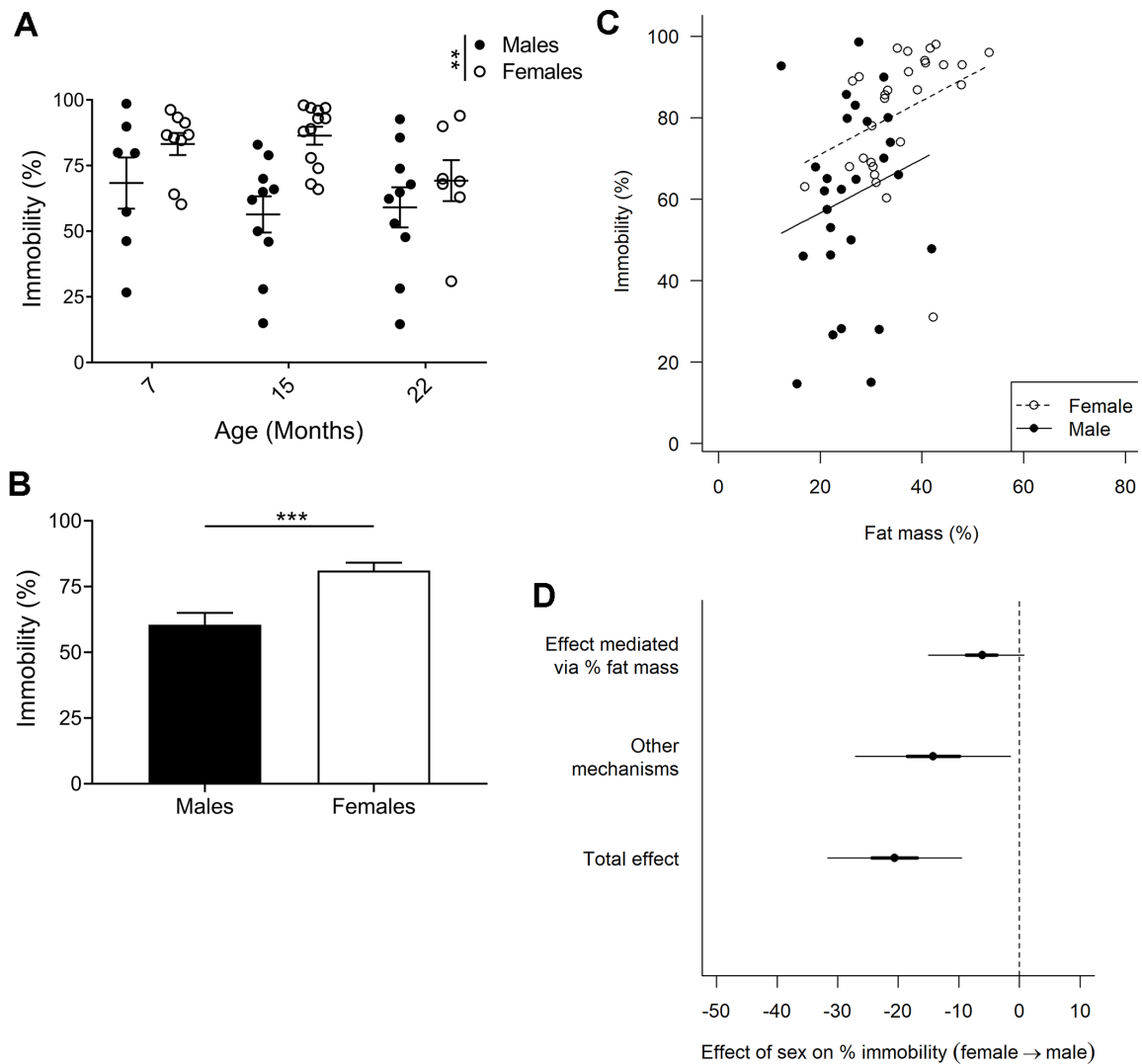


**Figure 4. Learning and memory assessed by shuttle box passive avoidance test.** Data are presented as the time it took to enter the conditioned area on the testing day (Day 2) subtracted by the time to enter on the conditioning day (Day 1), calculated for each individual animal, and indicate no major age or sex difference ( $p_{M7-F7}=0.055$ ,  $p_{M15-F15}=0.97$ ,  $p_{M22-F22}=0.74$ ,  $p_{M7-15}=0.11$ ,  $p_{M7-22}=0.23$ ,  $p_{M15-22}=0.19$ ,  $p_{F7-15}=0.62$ ,  $p_{F7-22}=0.30$ ,  $p_{F15-22}=0.60$ ; Mann-Whitney). Values are mean  $\pm$  SEM;  $n_{M7}=7$ ,  $n_{M15}=10$ ,  $n_{M22}=10$ ,  $n_{F7}=9$ ,  $n_{F15}=12$ ,  $n_{F22}=7$ .

gives a qualitative estimation of the natural lifespan without the expense of animal suffering.

In our lifespan study, mice were euthanized upon signs of pain or severe disease. Counting these animals as diseased by natural cause generates an underestimation of natural lifespan, generally defined as lifetime until animals die from intrinsic causes. When euthanized animals were instead "censored" in the data analysis (i.e. counted as healthy when terminated instead of

diseased), natural lifespan is overestimated since the removed animals were suffering from severe pathologies that likely limit their life expectancy. Thus, the natural lifespan is somewhere in between these lifespan calculations. Using this estimation, we categorized the 7-month-old mice as mature healthy adults with 100% survival, the 15-month-old timepoint as middle-aged with approximately 90% survival, and the 22-month-old mice as old with around 71% survival for males and 74% survival for females.



**Figure 5. Forced swim test to assay depressive-like behavior.** (A) Immobility in the forced swim test is considered an indicator of depressive-like behavior, and F2 hybrid females were found to be more immobile ( $p=0.0013$ , 2-way ANOVA), as compared to males. Values are mean  $\pm$  SEM. (B) Due to no age effect found, all timepoints for each sex were pooled to demonstrate a 20% difference in immobility between females and males ( $p=0.0002$ , Mann-Whitney). Values are mean  $\pm$  SEM. (C) Immobility plotted against fat mass for each animal. Females exhibited increased fat mass as compared to males, and immobility was found to correlate to percent fat mass. The correlation slopes were constrained to be equal since they were not significantly different ( $p=0.58$ , ANCOVA). (D) Bayesian mediation analysis separated the total effect of the difference between males and females in the forced swim test into effect mediated by fat mass or by unknown mechanisms. The sex difference in fat mass accounts for 30% of the sex difference in immobility (Bayesian  $p=0.96$ ) and other mechanisms account for the remaining 70% (Bayesian  $p=0.99$ ; Bayesian mediation analysis). Values are mean  $\pm$  50% (thick lines) and 95% confidence interval (thin lines);  $n_{M7}=7$ ,  $n_{M15}=10$ ,  $n_{M22}=10$ ,  $n_{F7}=9$ ,  $n_{F15}=12$ ,  $n_{F22}=7$ .

The interval of median lifespan of our C57BL/6N×BALB/c F2 male mice was in the same range as males of the four-way cross UM-HET3 mice, progeny of BALB/cByJ×C57BL/6J F1 females and C3H/HeJ×DBA/2J F1 males (789±130 compared to 742-826 days) and just below that for females (801±88 compared to 832-891 days) [51]. Compared to inbred mice, the median lifespan of both male and female C57BL/6N×BALB/c F2 mice fall in between that of the short-lived BALB/c and the long-lived C57BL/6J in the Aging Phenome Project (711/901 days for males and 771/866 days for females) [52], in which the Aging Center at the Jackson Laboratory examined the lifespan of 32 inbred mouse strains. Based on the same updated dataset [53], onset of death and maximum lifespan of our F2 males were similar to that of BALB/c males, while our F2 females more closely resembled C57BL/6J females, with similar onset of death. Additionally, only one C57BL/6J female outlived the longest-lived F2 hybrid female (the 3 longest-lived C57BL/6J females out of the 29 natural deaths died at 1180/1049/1049 days, compared to the 3 longest-lived hybrid mice out of the 13 natural deaths that died at 1051/1001/966 days).

In this study, we demonstrate that exploratory activity decreased as hybrid mice aged from 7 to 22 months. Although an age-related decline in exploratory behavior has been reported in several inbred strains, our findings confirm the pervasiveness of an age-related gradual decrease in exploratory movement for both females and males in heterogeneous hybrid mice. We thus herein have identified a decrease in exploratory behavior as a conserved non-biased behavioral hallmark of aging. Raising awareness of explorative activity in mice is vastly important in behavioral studies since the drive to explore affects fundamental incentives and can alter behavior in other tests that rely on exploration [23, 24].

In our F2 hybrid study, the mice displayed a relatively stable body weight in all age groups; likely because the first timepoint truly represented “mature adults”, and the last “old” age group represented a healthily aged cohort, thanks to the strict animal ethics protocol. Although the average body weight was comparable, female mice had more fat mass and less lean mass as compared to males, at least up to 15 months of age. In line with our results, similar body weights with increased fat mass and reduced lean mass in females as compared to males were previously observed in 20-month-old F1 hybrid C57BL/6J×129S1/SvImJ mice [54]. Additionally, body composition comparisons at 4 months of age of inbred mice demonstrated that males tended to have more lean mass, though also higher body weight, in more than half of the 40 different strains analyzed [55].

Results from this study pinpoint the importance to examine body composition in water-based behavioral studies. Fat mass content correlated to immobility in the forced swim test, which measures depressive-like behavior, with the elevated fat mass in females versus males directly accounting for 30–46% of the observed increase in depressive-like behavior. Notably, buoyancy caused by trapped air in the fur has previously been linked to swimming immobility [56], although the dependence of body fat mass in the forced swim test performance has not been identified before. We hypothesize that the increased floating sensation applied by decreased body density impacts the incentive to move. If so, fat mass likely impacts mouse activity, capacity, and behavior in water. Generally, only body weight is reported in behavioral analyses, and conclusions of depressive-like behavior using the forced swim test in obesity models, different strains, and sex comparisons have been drawn without taking body composition into account. In our study, fat mass directly accounted for 30–46% of the sex difference observed in the forced swim test, leaving 54–70% to be explained by unknown mechanisms. These may, in turn, be indirectly linked to fat mass through physiological effects. One example of this could be a reduced contractile function of muscle reported in obese mice [reviewed in 57], a finding that may also restrict water mobility in females. There are a number of potential fat mass unrelated mechanisms that could also play a role, including sex hormone levels as well as hormone signaling and their effects on neuronal circuits (though both of these examples are known to interact with metabolism). Taken together, we can only speculate and cannot exclude that females exhibited enhanced depressive-like behavior regardless of body composition. However, the 30-46% effect due to fat mass could still be the single most important factor on increased depressive-like behavior in females as compared to males. In addition, we did not identify any mobility alterations upon aging in the forced swim test in either males or females.

We found no significant sex or age effect on memory of passive avoidance learning using a one-time pain-conditioning factor. To our knowledge, only Benice and colleagues [30] have investigated aging in female and male mice using a passive avoidance test prior to us. In that study, inbred mice were conditioned the first day until they did not enter the dark compartment for three consecutive trials with a maximum of 10 trials [30]. Although it was stated that there were no significant age or sex differences regarding the conditioning trials, the cognitive learning and memory responses after multiple electric shocks enforces amygdala over hippocampal dependency of the test and increases memory consolidation [58, 59], and thus convolutes



possible comparisons and extrapolations to our study. Nonetheless, our data are in line with active avoidance and object memory consolidation tests of learning and memory in inbred mice, in which females and males performed equally well [10, 23, 30, 35]. We hypothesize that different age-related cognitive impairments are reflected depending on the test used. In order to resolve any controversies surrounding sex differences in various learning and memory tests, further analyses are needed to ensure that spatial testing is not sex-biased for reference cues. Moreover, we also strongly recommend that body composition and buoyancy are taken into account when analyzing results from water-based cognition tests, including the commonly used Morris water maze.

In summary, this work is the first behavioral phenotypic aging study to use hybrid mice and include analyses of both sexes. We herein have confirmed that decreased exploratory behavior is a conserved robust marker of aging, while no observable differences in general activity or in memory of passive avoidance learning were found. We have also demonstrated that increased fat mass partly explains why females swim less than males in the forced swim test of depressive-like behavior. This novel finding emphasizes the need to control for body composition in water-based tests. Moreover, we present a new method to qualitatively estimate natural lifespan in which animals are euthanized upon pain or severe disease.

## METHODS

### Animals and diets

Subjects were male and female C57BL/6N×BALB/c F2 hybrid mice crossed from C57BL/6N×BALB/c F1 hybrids, with C57BL/6N fathers (Charles River, Lyon, France) and BALB/c mothers (Harlan Laboratories, Horst, the Netherlands), with coat color ranging from white to brown and black. Males were cohoused in groups of three until 6 months of age. In 11 cages with male mice, fighting led to severe injury and all animals in those cages were removed from the study. This was done in order to maintain a representative population, since high-ranked animals would otherwise be selected for. All males were single-housed for the remaining time from 6 months of age. Females were cohoused in groups of four, with the number of animals per cage decreasing as mice deceased in the lifespan study. The mice were not pooled. Cages were equipped with nesting material (paper), cardboard houses, and wooden sticks, and were cleaned every two weeks for single housed males and every week for cohoused mice. Old nesting material was transferred to the new cages upon cleaning. The mice had *ad libitum* access to regular chow diet (R3; Lactamin, Kimstad Sweden) containing 12% fat, 62% carbohydrates, and

26% protein (energy percentage), with a total energy content of 3 Kcal/g. The mice were maintained on a 12:12 h light/dark cycle (gradual light increase from 5:30-6:00 am and decrease at 5:30-6:00 pm) at 21°C. All experimental protocols of the study were carried out in accordance with the ethical certificate approved by the Animal Ethics Committee in Gothenburg, Sweden (Permit Number: 164-2015) and with EU Directive 2010/63/EU for animal experiments. Causes of euthanization were symptoms indicative of severe ill-health. These include: hunched shoulders, shabby fur, inactivity, failure to eat or drink, enlarged organs and tumors. Upon termination, mice were euthanized by decapitation under 5% isoflurane anesthesia.

### Study design of physiological and behavioral phenotypic profiling

The lifespan study consisted of 20 male and 22 female mice, which were not part of the behavioral testing cohorts. Littermates to these mice were tested for physiological and behavioral phenotypic profiling at three ages representing mature adults (7 months; 7 males and 10 females), middle-aged (15 months; 10 males and 12 females) and aged (22 months; 10 males and 7 females) [60]. Specifically, the mice's ages at behavioral profiling were  $6.6-7.8 \pm 0.2$  (age at test period start - age at test period end  $\pm$ SD of age variance in the cohort; 7 months),  $14.5-15.6 \pm 0.1$  (15 months),  $21.8-22.5 \pm 0.2$  (22 months). These mice were initially included in the lifespan analysis but taken out (marked as censored) at the testing timepoint, and thus were only included in the calculation of percent survival in the lifespan analysis until their removal. Results from the 7-month cohorts served as control in a previous study [61]. Test periods occurred at different times of the year without any attempt to control for seasonal effects (7-month-olds: September-October; 15-month-olds: May-June; 22-month-olds: late December - January). Prior to the testing period, all mice were handled and acclimated for one week. The sequence of analyses is provided in Supplementary Table 3. Due to technical and practical reasons in our animal facility, the order of the behavioral tests could not be identical for males and females for all timepoints. Behavioral tests that are known to be impacted by handling (such as activity box) were performed early in the testing period (except for the 15-month-old female activity box timepoint; conclusions from which are therefore omitted). No correlation between behavioral test results and coat color could be found. Tests that affect the mice to a greater extent were planned later in the testing periods and/or the animals were given an extended recovery time in order to not impact the subsequent analyses. All animals were acclimated to the behavioral testing room for at least one hour prior to testing, and all experiments started between 10-11 am with the time spent for each experiment noted

for each individual test. One female in the 7-month timepoint was removed from testing due to an eye injury.

### **Body composition and core temperature measurements**

Core body temperature, body weight, and body composition were obtained at the same time. Core body temperature was taken with a rectal probe thermometer (ELFA AB, Sweden) followed by dual energy X-ray absorptiometry (DEXA) scanning (Lunar PIXImus Densitometer, GE Medical Systems, Madison, WI, USA) while under 2.5% isoflurane sedation for approximately 4 minutes. The parameters recorded were: body length (cm), body fat (g), lean body mass (g), bone mineral density (BMD; g/cm<sup>2</sup>), and bone mineral content (BMC; g) [47].

### **Activity box**

The activity box is an open-field activity-like test to assess general activity, exploratory behavior, and signs of anxiety [47]. The activity box records movement of each mouse three dimensionally using infrared sensors built into the walls (8Lx8Bx8H) of a sound-proof opaque box (50x50x50 cm) with a low intensity lamp over the lid (Kungsbacka mät och regler, Fjärrås, Sweden). The mice were placed in the middle of the box and their movements were recorded for a total of 1 hour. The first 5 minutes were used to assess exploratory behavior in this novel environment and the full hour was used to assess general activity and behavior. The parameters that were recorded as events every 5 minutes include: horizontal activity, peripheral activity, rearing activity, peripheral rearing, rearing time, locomotion, and corner time.

### **Forced swim test**

The forced swim test is used to analyze mice for signs of depression [47, 62, 63]. A transparent plexiglas cylinder (25 cm inner-diameter, 60 cm length) was filled with room temperature (22 °C) water level with a grey circular plastic platform hanging from wires on the outside of the cylinder, approximately 20 cm from the top (bespoke construction, AstraZeneca, Gothenburg, Sweden). A single mouse was placed on the water surface and monitored for 6 minutes and 20 seconds by a video camera placed directly above the cylinder, with the last 4 minutes used to analyze behavior and activity (MouseTracker analysis software, Mölndal, Sweden).

### **Shuttle box passive avoidance test**

The shuttle box passive avoidance test is used to study memory performance in mice [47, 64]. The shuttle box

system (Accuscan Instruments Inc., Columbus, OH, USA) is made up of a cage centrally divided by a wall with a sliding door creating two compartments, one bright with transparent walls and one dark covered on all sides with opaque walls. Both compartments are equipped with sensors that determine the location of the mouse, and the floor of the cage is made of a stainless-steel grid that can deliver a mild electric shock. On the first day, each mouse was placed into the bright compartment, and after 60 seconds the central door opened allowing the mouse to migrate into the dark compartment, an environment which they should prefer. Upon entry into the dark compartment, immediately after the door closed, the mouse received a mild electric shock (0.3 mA) and remained in the dark compartment for at least 30 seconds thereafter. On the second day, the mouse was again placed into the bright compartment and after 60 seconds when the central door opened the mouse was given the same choice to enter into the dark compartment (300 seconds maximum time). The time of entry into the dark compartment was recorded for both days, and no entry or entry after a longer interval on the second day as compared to the first day was considered a memory response.

### **Statistical analysis**

Comparisons between two groups were analyzed for normal distribution by Shapiro-Wilks and Levene's test for homogeneity of variance. Two-tailed independent t-test (Student's t-test) was used for groups that met these criteria, two-tailed independent unequal variance t-test (Welch's test) for normal distributed groups which failed Levene's test, and nonparametric Mann-Whitney U test for unevenly distributed groups. Comparisons of survival curves were analyzed with the log-rank Mantel-Cox test. Interactions and comparisons of age and sex effects were calculated with 2-way ANOVA. Covariance was analyzed with ANCOVA. Differences were considered significant at  $P < 0.05$ . A Bayesian mediation analysis was used to test whether sex differences on the forced swim test were due to difference in fat mass between sexes, or due to other mechanisms [described in 49, 50]. The Bayesian p-value reported in the results has a direct and intuitive interpretation: it is the probability of the conclusion given the data, so a high value means that the conclusion is highly probable. Statistics were calculated using IBM SPSS Statistics 25, GraphPad Prism 7, or R/Stan.

### **ACKNOWLEDGMENTS**

We thank Liselotte Andersson for excellent animal husbandry; Richard Miller for valuable input on experimental design on mouse lifespan studies; Mikael

Bjursell and Johan K. Johansson for helpful contribution to the phenotypic study setup; Johan K. Johansson and Michelle Porritt for technical support with behavioral testing and constructive discussions in regard to the tests.

## CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interests.

## FUNDING

This work was financially supported by the Swedish Foundation for Strategic Research (SSF; ID14-0087), AstraZeneca AB, The National Institute on Aging (K99AG05583 to JMR) and Konung Gustav V:s och Drottning Victorias Frimurarestiftelse (JMR). None of these organizations had any role in study design, data analysis, data interpretation, or manuscript writing.

## REFERENCES

1. Lipman RD. Pathobiology of aging rodents: inbred and hybrid models. *Exp Gerontol*. 1997; 32:215–28.  
[https://doi.org/10.1016/S0531-5565\(96\)00037-X](https://doi.org/10.1016/S0531-5565(96)00037-X)  
PMID:[9088918](https://pubmed.ncbi.nlm.nih.gov/9088918/)
2. Crawley JN, Belknap JK, Collins A, Crabbe JC, Frankel W, Henderson N, Hitzemann RJ, Maxson SC, Miner LL, Silva AJ, Wehner JM, Wynshaw-Boris A, Paylor R. Behavioral phenotypes of inbred mouse strains: implications and recommendations for molecular studies. *Psychopharmacology (Berl)*. 1997; 132:107–24.  
<https://doi.org/10.1007/s002130050327>  
PMID:[9266608](https://pubmed.ncbi.nlm.nih.gov/9266608/)
3. Vöikar V, Vasar E, Rauvala H. Behavioral alterations induced by repeated testing in C57BL/6J and 129S2/Sv mice: implications for phenotyping screens. *Genes Brain Behav*. 2004; 3:27–38.  
<https://doi.org/10.1046/j.1601-183X.2003.0044.x>  
PMID:[14960013](https://pubmed.ncbi.nlm.nih.gov/14960013/)
4. Miller RA, Nadon NL. Principles of Animal Use for Gerontological Research. *J Gerontol A Biol Sci Med Sci*. 2000; 55:117–123.  
<https://doi.org/10.1093/gerona/55.3.B117>  
PMID:[10795715](https://pubmed.ncbi.nlm.nih.gov/10795715/)
5. Miller RA, Austad S, Burke D, Chrisp C, Dysko R, Galecki A, Jackson A, Monnier V. Exotic mice as models for aging research: polemic and prospectus. *Neurobiol Aging*. 1999; 20:217–31.  
PMID:[10537031](https://pubmed.ncbi.nlm.nih.gov/10537031/)
6. Sumien N, Sims MN, Taylor HJ, Forster MJ. Profiling psychomotor and cognitive aging in four-way cross mice. *Age (Dordr)*. 2006; 28:265–82.  
<https://doi.org/10.1007/s11357-006-9015-7>  
PMID:[22253494](https://pubmed.ncbi.nlm.nih.gov/22253494/)
7. Rivera J, Tessarollo L. Genetic background and the dilemma of translating mouse studies to humans. *Immunity*. 2008; 28:1–4.  
<https://doi.org/10.1016/j.immuni.2007.12.008>  
PMID:[18199409](https://pubmed.ncbi.nlm.nih.gov/18199409/)
8. Beery AK, Zucker I. Sex bias in neuroscience and biomedical research. *Neurosci Biobehav Rev*. 2011; 35:565–72.  
<https://doi.org/10.1016/j.neubiorev.2010.07.002>  
PMID:[20620164](https://pubmed.ncbi.nlm.nih.gov/20620164/)
9. Yoon DY, Mansukhani NA, Stubbs VC, Helenowski IB, Woodruff TK, Kibbe MR. Sex bias exists in basic science and translational surgical research. *Surgery*. 2014; 156:508–16.  
<https://doi.org/10.1016/j.surg.2014.07.001>  
PMID:[25175501](https://pubmed.ncbi.nlm.nih.gov/25175501/)
10. Frick KM, Burlingame LA, Arters JA, Berger-Sweeney J. Reference memory, anxiety and estrous cyclicity in C57BL/6NIA mice are affected by age and sex. *Neuroscience*. 2000; 95:293–307.  
[https://doi.org/10.1016/S0306-4522\(99\)00418-2](https://doi.org/10.1016/S0306-4522(99)00418-2)  
PMID:[10619486](https://pubmed.ncbi.nlm.nih.gov/10619486/)
11. Kelly SJ, Ostrowski NL, Wilson MA. Gender differences in brain and behavior: hormonal and neural bases. *Pharmacol Biochem Behav*. 1999; 64:655–64.  
[https://doi.org/10.1016/S0091-3057\(99\)00167-7](https://doi.org/10.1016/S0091-3057(99)00167-7)  
PMID:[10593187](https://pubmed.ncbi.nlm.nih.gov/10593187/)
12. Vöikar V, Köks S, Vasar E, Rauvala H. Strain and gender differences in the behavior of mouse lines commonly used in transgenic studies. *Physiol Behav*. 2001; 72:271–81.  
[https://doi.org/10.1016/S0031-9384\(00\)00405-4](https://doi.org/10.1016/S0031-9384(00)00405-4)  
PMID:[11240006](https://pubmed.ncbi.nlm.nih.gov/11240006/)
13. Bekker MH, van Mens-Verhulst J. Anxiety disorders: sex differences in prevalence, degree, and background, but gender-neutral treatment. *Gend Med*. 2007 (Suppl B); 4:S178–93.  
[https://doi.org/10.1016/S1550-8579\(07\)80057-X](https://doi.org/10.1016/S1550-8579(07)80057-X)  
PMID:[18156102](https://pubmed.ncbi.nlm.nih.gov/18156102/)
14. Palanza P. Animal models of anxiety and depression: how are females different? *Neurosci Biobehav Rev*. 2001; 25:219–33.  
[https://doi.org/10.1016/S0149-7634\(01\)00010-0](https://doi.org/10.1016/S0149-7634(01)00010-0)  
PMID:[11378178](https://pubmed.ncbi.nlm.nih.gov/11378178/)
15. Frackiewicz EJ, Sramek JJ, Cutler NR. Gender differences in depression and antidepressant pharmacokinetics and adverse events. *Ann Pharmacother*. 2000; 34:80–88.  
<https://doi.org/10.1345/aph.18465>  
PMID:[10669189](https://pubmed.ncbi.nlm.nih.gov/10669189/)

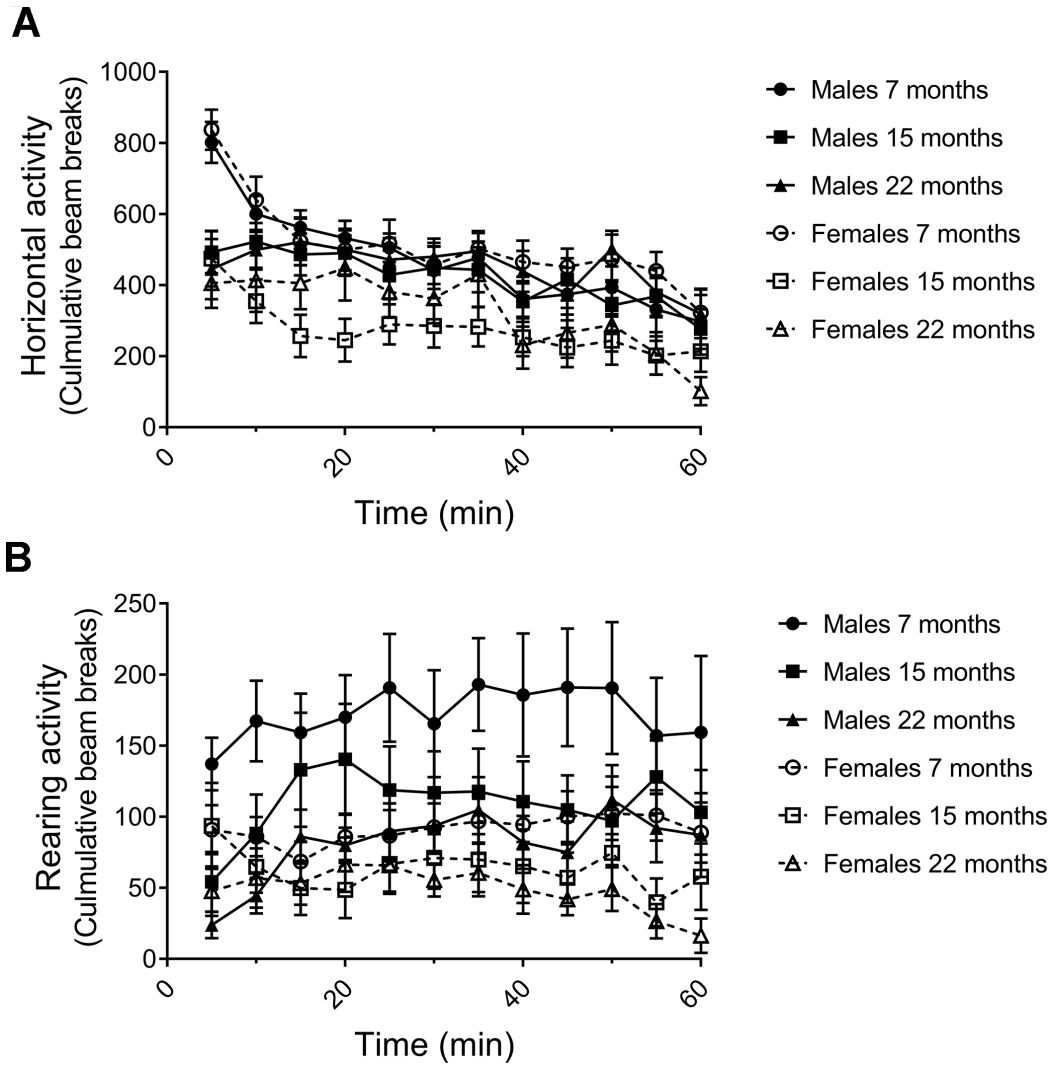
16. Gorman JM. Gender differences in depression and response to psychotropic medication. *Gend Med.* 2006; 3:93–109.  
[https://doi.org/10.1016/S1550-8579\(06\)80199-3](https://doi.org/10.1016/S1550-8579(06)80199-3)  
PMID:[16860269](https://pubmed.ncbi.nlm.nih.gov/16860269/)
17. Haack S, Seeringer A, Thürmann P, Becker T, Kirchheiner J. Sex-specific differences in side effects of psychotropic drugs: Genes or gender? *Pharmacogenomics.* 2009; 10:1511–26.  
<https://doi.org/10.2217/pgs.09.102> PMID:[19761372](https://pubmed.ncbi.nlm.nih.gov/19761372/)
18. Meziane H, Ouagazzal AM, Aubert L, Wietrzych M, Krezel W. Estrous cycle effects on behavior of C57BL/6J and BALB/cByJ female mice: implications for phenotyping strategies. *Genes Brain Behav.* 2007; 6:192–200.  
<https://doi.org/10.1111/j.1601-183X.2006.00249.x>  
PMID:[16827921](https://pubmed.ncbi.nlm.nih.gov/16827921/)
19. Palanza P, Gioiosa L, Parmigiani S. Social stress in mice: gender differences and effects of estrous cycle and social dominance. *Physiol Behav.* 2001; 73:411–20.  
[https://doi.org/10.1016/S0031-9384\(01\)00494-2](https://doi.org/10.1016/S0031-9384(01)00494-2)  
PMID:[11438369](https://pubmed.ncbi.nlm.nih.gov/11438369/)
20. Crawley JN. Behavioral phenotyping of transgenic and knockout mice: experimental design and evaluation of general health, sensory functions, motor abilities, and specific behavioral tests. *Brain Res.* 1999; 835:18–26.  
[https://doi.org/10.1016/S0006-8993\(98\)01258-X](https://doi.org/10.1016/S0006-8993(98)01258-X)  
PMID:[10448192](https://pubmed.ncbi.nlm.nih.gov/10448192/)
21. Crawley JN, Paylor R. A proposed test battery and constellations of specific behavioral paradigms to investigate the behavioral phenotypes of transgenic and knockout mice. *Horm Behav.* 1997; 31:197–211.  
<https://doi.org/10.1006/hbeh.1997.1382>  
PMID:[9213134](https://pubmed.ncbi.nlm.nih.gov/9213134/)
22. Rogers DC, Peters J, Martin JE, Ball S, Nicholson SJ, Witherden AS, Hafezparast M, Latcham J, Robinson TL, Quilter CA, Fisher EM. SHIRPA, a protocol for behavioral assessment: validation for longitudinal study of neurological dysfunction in mice. *Neurosci Lett.* 2001; 306:89–92.  
[https://doi.org/10.1016/S0304-3940\(01\)01885-7](https://doi.org/10.1016/S0304-3940(01)01885-7)  
PMID:[11403965](https://pubmed.ncbi.nlm.nih.gov/11403965/)
23. Fahlström A, Zeberg H, Ulfhake B. Changes in behaviors of male C57BL/6J mice across adult life span and effects of dietary restriction. *Age (Dordr).* 2012; 34:1435–52.  
<https://doi.org/10.1007/s11357-011-9320-7>  
PMID:[21989972](https://pubmed.ncbi.nlm.nih.gov/21989972/)
24. Fahlström A, Yu Q, Ulfhake B. Behavioral changes in aging female C57BL/6 mice. *Neurobiol Aging.* 2011; 32:1868–80.  
<https://doi.org/10.1016/j.neurobiolaging.2009.11.003>  
PMID:[20005598](https://pubmed.ncbi.nlm.nih.gov/20005598/)
25. Dean RL 3rd, Scozzafava J, Goas JA, Regan B, Beer B, Bartus RT 3rd. Age-related differences in behavior across the life span of the C57BL/6J mouse. *Exp Aging Res.* 1981; 7:427–51.  
<https://doi.org/10.1080/03610738108259823>  
PMID:[7333338](https://pubmed.ncbi.nlm.nih.gov/7333338/)
26. Shoji H, Takao K, Hattori S, Miyakawa T. Age-related changes in behavior in C57BL/6J mice from young adulthood to middle age. *Mol Brain.* 2016; 9:11.  
<https://doi.org/10.1186/s13041-016-0191-9>  
PMID:[26822304](https://pubmed.ncbi.nlm.nih.gov/26822304/)
27. Jonasson Z. Meta-analysis of sex differences in rodent models of learning and memory: a review of behavioral and biological data. *Neurosci Biobehav Rev.* 2005; 28:811–25.  
<https://doi.org/10.1016/j.neubiorev.2004.10.006>  
PMID:[15642623](https://pubmed.ncbi.nlm.nih.gov/15642623/)
28. Koss WA, Frick KM. Sex differences in hippocampal function. *J Neurosci Res.* 2017; 95:539–62.  
<https://doi.org/10.1002/jnr.23864> PMID:[27870401](https://pubmed.ncbi.nlm.nih.gov/27870401/)
29. Foster TC, Defazio RA, Bizon JL. Characterizing cognitive aging of spatial and contextual memory in animal models. *Front Aging Neurosci.* 2012; 4:12–12.  
<https://doi.org/10.3389/fnagi.2012.00012>  
PMID:[22988436](https://pubmed.ncbi.nlm.nih.gov/22988436/)
30. Benice TS, Rizk A, Kohama S, Pfankuch T, Raber J. Sex-differences in age-related cognitive decline in C57BL/6J mice associated with increased brain microtubule-associated protein 2 and synaptophysin immunoreactivity. *Neuroscience.* 2006; 137:413–23.  
<https://doi.org/10.1016/j.neuroscience.2005.08.029>  
PMID:[16330151](https://pubmed.ncbi.nlm.nih.gov/16330151/)
31. de Fiebre NC, Sumien N, Forster MJ, de Fiebre CM. Spatial learning and psychomotor performance of C57BL/6 mice: age sensitivity and reliability of individual differences. *Age (Dordr).* 2006; 28:235–53.  
<https://doi.org/10.1007/s11357-006-9027-3>  
PMID:[22253492](https://pubmed.ncbi.nlm.nih.gov/22253492/)
32. Frick KM, Gresack JE. Sex differences in the behavioral response to spatial and object novelty in adult C57BL/6 mice. *Behav Neurosci.* 2003; 117:1283–91.  
<https://doi.org/10.1037/0735-7044.117.6.1283>  
PMID:[14674847](https://pubmed.ncbi.nlm.nih.gov/14674847/)
33. Berger-Sweeney J, Arnold A, Gabeau D, Mills J. Sex differences in learning and memory in mice: effects of sequence of testing and cholinergic blockade. *Behav Neurosci.* 1995; 109:859–73.  
<https://doi.org/10.1037/0735-7044.109.5.859>  
PMID:[8554711](https://pubmed.ncbi.nlm.nih.gov/8554711/)

34. LaBuda CJ, Mellgren RL, Hale RL. Sex differences in the acquisition of a radial maze task in the CD-1 mouse. *Physiol Behav.* 2002; 76:213–17.  
[https://doi.org/10.1016/S0031-9384\(02\)00713-8](https://doi.org/10.1016/S0031-9384(02)00713-8)  
PMID:[12044593](https://pubmed.ncbi.nlm.nih.gov/12044593/)
35. Mishima N, Higashitani F, Teraoka K, Yoshioka R. Sex differences in appetitive learning of mice. *Physiol Behav.* 1986; 37:263–68.  
[https://doi.org/10.1016/0031-9384\(86\)90230-1](https://doi.org/10.1016/0031-9384(86)90230-1)  
PMID:[3737737](https://pubmed.ncbi.nlm.nih.gov/3737737/)
36. Sandstrom NJ, Kaufman J, Huettel SA. Males and females use different distal cues in a virtual environment navigation task. *Brain Res Cogn Brain Res.* 1998; 6:351–60.  
[https://doi.org/10.1016/S0926-6410\(98\)00002-0](https://doi.org/10.1016/S0926-6410(98)00002-0)  
PMID:[9593991](https://pubmed.ncbi.nlm.nih.gov/9593991/)
37. Kant L, Yilmaz O, Taskiran D, Kulali B, Furedy JJ, Demirgören S, Pöğün S. Sexually dimorphic cognitive style, female sex hormones, and cortical nitric oxide. *Physiol Behav.* 2000; 71:277–87.  
[https://doi.org/10.1016/S0031-9384\(00\)00327-9](https://doi.org/10.1016/S0031-9384(00)00327-9)  
PMID:[11150559](https://pubmed.ncbi.nlm.nih.gov/11150559/)
38. Roof RL, Stein DG. Gender differences in Morris water maze performance depend on task parameters. *Physiol Behav.* 1999; 68:81–86.  
[https://doi.org/10.1016/S0031-9384\(99\)00162-6](https://doi.org/10.1016/S0031-9384(99)00162-6)  
PMID:[10627065](https://pubmed.ncbi.nlm.nih.gov/10627065/)
39. Kanit L, Taşkıran D, Furedy JJ, Kulali B, McDonald R, Pöğün S. Nicotine interacts with sex in affecting rat choice between “look-out” and “navigational” cognitive styles in the Morris water maze place learning task. *Brain Res Bull.* 1998; 46:441–45.  
[https://doi.org/10.1016/S0361-9230\(98\)00008-2](https://doi.org/10.1016/S0361-9230(98)00008-2)  
PMID:[9739007](https://pubmed.ncbi.nlm.nih.gov/9739007/)
40. David DJ, Renard CE, Jolliet P, Hascoët M, Bourin M. Antidepressant-like effects in various mice strains in the forced swimming test. *Psychopharmacology (Berl).* 2003; 166:373–82.  
<https://doi.org/10.1007/s00213-002-1335-4>  
PMID:[12601501](https://pubmed.ncbi.nlm.nih.gov/12601501/)
41. Fernández-Guasti A, Olivares-Nazario M, Reyes R, Martínez-Mota L. Sex and age differences in the antidepressant-like effect of fluoxetine in the forced swim test. *Pharmacol Biochem Behav.* 2017; 152:81–89.  
<https://doi.org/10.1016/j.pbb.2016.01.011>  
PMID:[26807812](https://pubmed.ncbi.nlm.nih.gov/26807812/)
42. Bourin M, Colombel MC, Redrobe JP, Nizard J, Hascoët M, Baker GB. Evaluation of efficacies of different classes of antidepressants in the forced swimming test in mice at different ages. *Prog Neuropsychopharmacol Biol Psychiatry.* 1998; 22:343–51.  
[https://doi.org/10.1016/S0278-5846\(98\)00009-8](https://doi.org/10.1016/S0278-5846(98)00009-8)  
PMID:[9608606](https://pubmed.ncbi.nlm.nih.gov/9608606/)
43. Porsolt RD, Bertin A, Jalfre M. “Behavioural despair” in rats and mice: strain differences and the effects of imipramine. *Eur J Pharmacol.* 1978; 51:291–94.  
[https://doi.org/10.1016/0014-2999\(78\)90414-4](https://doi.org/10.1016/0014-2999(78)90414-4)  
PMID:[568552](https://pubmed.ncbi.nlm.nih.gov/568552/)
44. Bogdanova OV, Kanekar S, D’Anci KE, Renshaw PF. Factors influencing behavior in the forced swim test. *Physiol Behav.* 2013; 118:227–39.  
<https://doi.org/10.1016/j.physbeh.2013.05.012>  
PMID:[23685235](https://pubmed.ncbi.nlm.nih.gov/23685235/)
45. Lucki I, Dalvi A, Mayorga AJ. Sensitivity to the effects of pharmacologically selective antidepressants in different strains of mice. *Psychopharmacology (Berl).* 2001; 155:315–22.  
<https://doi.org/10.1007/s002130100694>  
PMID:[11432695](https://pubmed.ncbi.nlm.nih.gov/11432695/)
46. Tanaka S, Young JW, Halberstadt AL, Masten VL, Geyer MA. Four factors underlying mouse behavior in an open field. *Behav Brain Res.* 2012; 233:55–61.  
<https://doi.org/10.1016/j.bbr.2012.04.045>  
PMID:[22569582](https://pubmed.ncbi.nlm.nih.gov/22569582/)
47. Gerdin AK, Surve VV, Jönsson M, Bjursell M, Björkman M, Edenro A, Schuelke M, Saad A, Bjurström S, Lundgren EJ, Snaith M, Fransson-Steen R, Törnell J, et al. Phenotypic screening of hepatocyte nuclear factor (HNF) 4-γ receptor knockout mice. *Biochem Biophys Res Commun.* 2006; 349:825–32.  
<https://doi.org/10.1016/j.bbrc.2006.08.103>  
PMID:[16945327](https://pubmed.ncbi.nlm.nih.gov/16945327/)
48. Steenbergen HL, Heinsbroek RP, Van Hest A, Van de Poll NE. Sex-dependent effects of inescapable shock administration on shuttlebox-escape performance and elevated plus-maze behavior. *Physiol Behav.* 1990; 48:571–76.  
[https://doi.org/10.1016/0031-9384\(90\)90302-K](https://doi.org/10.1016/0031-9384(90)90302-K)  
PMID:[2075210](https://pubmed.ncbi.nlm.nih.gov/2075210/)
49. Lazic SE. Using causal models to distinguish between neurogenesis-dependent and -independent effects on behaviour. *J R Soc Interface.* 2012; 9:907–17.  
<https://doi.org/10.1098/rsif.2011.0510>  
PMID:[21957118](https://pubmed.ncbi.nlm.nih.gov/21957118/)
50. Lazic SE, Fuss J, Gass P. Quantifying the behavioural relevance of hippocampal neurogenesis. *PLoS One.* 2014; 9:e113855–113855.  
<https://doi.org/10.1371/journal.pone.0113855>  
PMID:[25426717](https://pubmed.ncbi.nlm.nih.gov/25426717/)
51. Strong R, Miller RA, Astle CM, Baur JA, de Cabo R, Fernandez E, Guo W, Javors M, Kirkland JL, Nelson JF, Sinclair DA, Teter B, Williams D, et al. Evaluation of resveratrol, green tea extract, curcumin, oxaloacetic

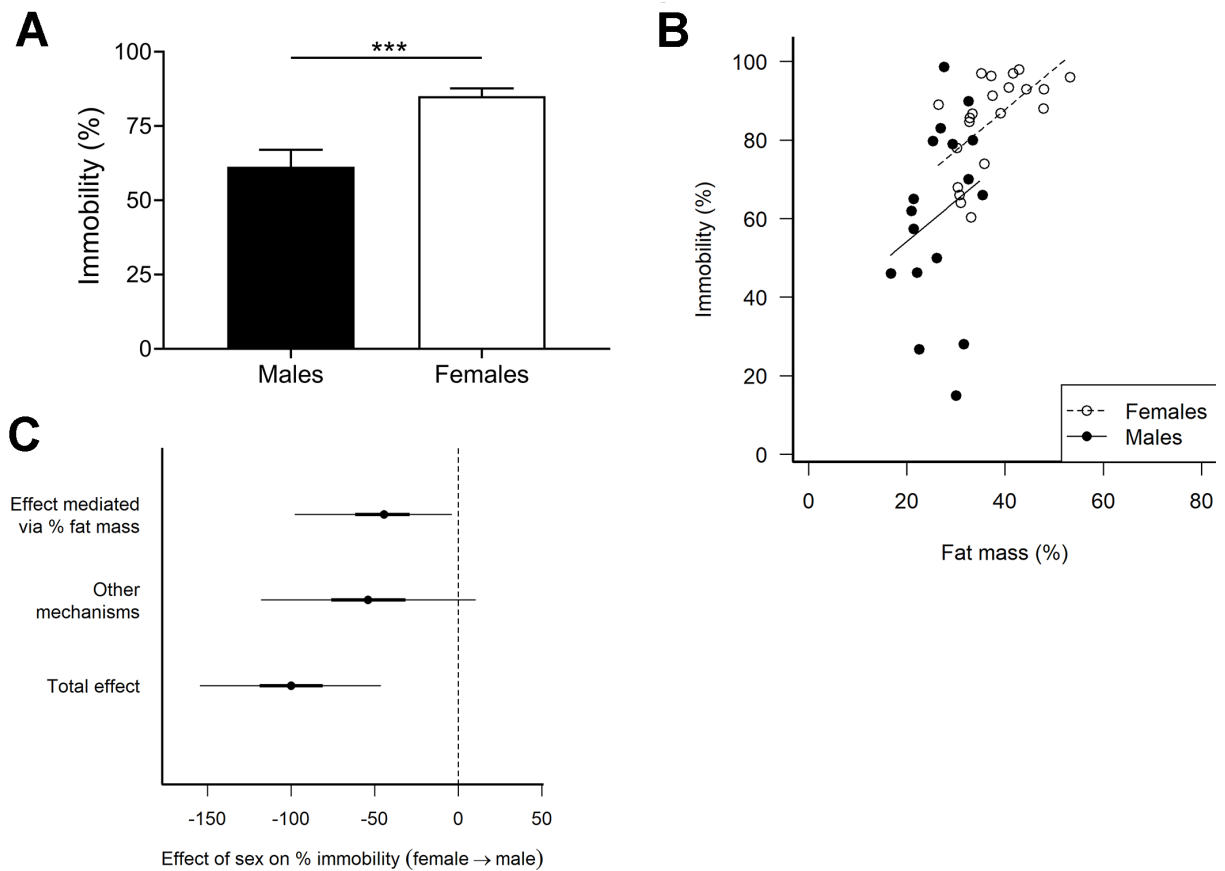
- acid, and medium-chain triglyceride oil on life span of genetically heterogeneous mice. *J Gerontol A Biol Sci Med Sci*. 2013; 68:6–16.  
<https://doi.org/10.1093/gerona/gls070>  
PMID:22451473
52. Yuan R, Tsaih SW, Petkova SB, Marin de Evsikova C, Xing S, Marion MA, Bogue MA, Mills KD, Peters LL, Bult CJ, Rosen CJ, Sundberg JP, Harrison DE, et al. Aging in inbred strains of mice: study design and interim report on median lifespans and circulating IGF1 levels. *Aging Cell*. 2009; 8:277–87.  
<https://doi.org/10.1111/j.1474-9726.2009.00478.x>  
PMID:19627267
53. Yuan R, Meng Q, Nautiyal J, Flurkey K, Tsaih SW, Krier R, Parker MG, Harrison DE, Paigen B. Genetic coregulation of age of female sexual maturation and lifespan through circulating IGF1 among inbred mouse strains. *Proc Natl Acad Sci USA*. 2012; 109:8224–29.  
<https://doi.org/10.1073/pnas.1121113109>  
PMID:22566614
54. Paigen B, Curren JM, Svenson KL. Effects of Varied Housing Density on a Hybrid Mouse Strain Followed for 20 Months. *PLoS One*. 2016; 11:e0149647.  
<https://doi.org/10.1371/journal.pone.0149647>  
PMID:26900840
55. Reed DR, Bachmanov AA, Tordoff MG. Forty mouse strain survey of body composition. *Physiol Behav*. 2007; 91:593–600.  
<https://doi.org/10.1016/j.physbeh.2007.03.026>  
PMID:17493645
56. Chen L, Faas GC, Ferando I, Mody I. Novel insights into the behavioral analysis of mice subjected to the forced-swim test. *Transl Psychiatry*. 2015; 5:e551.  
<https://doi.org/10.1038/tp.2015.44> PMID:25871976
57. Tallis J, James RS, Seebacher F. The effects of obesity on skeletal muscle contractile function. *J Exp Biol*. 2018; 221:jeb163840.  
<https://doi.org/10.1242/jeb.163840> PMID:29980597
58. Slotnick BM. Fear behavior and passive avoidance deficits in mice with amygdala lesions. *Physiol Behav*. 1973; 11:717–20.  
[https://doi.org/10.1016/0031-9384\(73\)90258-8](https://doi.org/10.1016/0031-9384(73)90258-8)  
PMID:4583907
59. Poulos AM, Mehta N, Lu B, Amir D, Livingston B, Santarelli A, Zhuravka I, Fanselow MS. Conditioning- and time-dependent increases in context fear and generalization. *Learn Mem*. 2016; 23:379–85.  
<https://doi.org/10.1101/lm.041400.115>  
PMID:27317198
60. Flurkey K, Curren J, Harrison DE. Chapter 20: Mouse Models in Aging Research. In: Fox JG, et al, editors. *The Mouse in Biomedical Research*. 2nd ed. Burlington: Academic Press; 2007. pp. 637–72.  
<https://doi.org/10.1016/B978-012369454-6/50074-1>
61. Adelöf J, Andersson M, Porritt M, Petersen A, Zetterberg M, Wiseman J, Hernebring M. PA28 $\alpha\beta$  overexpression enhances learning and memory of female mice without inducing 20S proteasome activity. *BMC Neurosci*. 2018; 19:70–70.  
<https://doi.org/10.1186/s12868-018-0468-2>  
PMID:30400847
62. Porsolt RD, Le Pichon M, Jalfre M. Depression: a new animal model sensitive to antidepressant treatments. *Nature*. 1977; 266:730–32.  
<https://doi.org/10.1038/266730a0> PMID:559941
63. Can A, Dao DT, Arad M, Terrillion CE, Piantadosi SC, Gould TD. The mouse forced swim test. *J Vis Exp*. 2012; 59:e3638.  
<https://doi.org/10.3791/3638> PMID:22314943
64. Bammer G. Pharmacological investigations of neurotransmitter involvement in passive avoidance responding: a review and some new results. *Neurosci Biobehav Rev*. 1982; 6:247–96.  
[https://doi.org/10.1016/0149-7634\(82\)90041-0](https://doi.org/10.1016/0149-7634(82)90041-0)  
PMID:6129605

SUPPLEMENTARY MATERIALS

Supplementary Figures



**Supplementary Figure 1. One-hour open field test activity in 5-minute bins.** (A) Horizontal activity declined intrasessionally for all cohorts, 7-month-old males and females demonstrated augmented activity in the initial phase of testing. (B) Rearing activity decreased with age but is maintained throughout the test hour regardless of age. Data represents mean  $\pm$  SEM.  $n_{M7}=7$ ,  $n_{M15}=10$ ,  $n_{M23}=10$ ,  $n_{F7}=10$ ,  $n_{F15}=12$ ,  $n_{F23}=7$



**Supplementary Figure 2. Immobility in the forced swim test for 7- and 15-month-old males and females.** (A) The sex difference in immobility when pooling females and males from the 7- and 15-month-old timepoints was 24% as compared to 21% when including 23 months animals ( $p=0.0007$ , Mann Whitney). Data represents mean  $\pm$  SEM. (B) Immobility plotted against fat mass for each animal from the 7 and 15 months' timepoints confirms the correlation of fat mass to immobility. (C) Since the sex difference in immobility remained similar (21–24%) but fat mass differences were most pronounced at 7 and 15 months, Bayesian mediation analysis was recalculated to answer how much of the effect was mediated by fat mass at 7 and 15 months. Results demonstrated that the effect mediated by fat mass increased to 46% ( $p=0.986$ ) of the total immobility effect and other unknown mechanisms accounted for 54% ( $p=0.952$ ) when older animals were excluded. This signifies that even though the sex difference is most likely independent of age, the effect mediated by fat mass is greater at younger ages (46% compared to 30%). Values are mean  $\pm$  50% (thick lines) and 95% confidence interval (thin lines);  $n_{M7}=7$ ,  $n_{M15}=10$ ,  $n_{F7}=10$ ,  $n_{F15}=12$ .



## Supplementary Tables

**Supplementary Table 1. Order of deaths in lifespan study.**

Age (days)	Animal	Cause of death	Age (days)	Animal	Cause of death
183	599, 2 (M), Brown	Timepoint 7 months	444	508, 5 (M), Brown	Timepoint 15 months
207	387, 1 (M), Black,	Timepoint 7 months	444	153, 1 (F), Brown	Timepoint 15 months
207	388, 1 (M), Brown	Timepoint 7 months	444	154, 2 (F), Brown	Timepoint 15 months
207	392, 5 (M), Brown	Timepoint 7 months	444	156, 2 (F), Albino	Timepoint 15 months
207	226, 2 (F), Brown	Timepoint 7 months	444	157, 3 (F), Albino	Timepoint 15 months
207	228, 4 (F), Brown	Timepoint 7 months	444	163, 3 (F), Brown	Timepoint 15 months
207	230, 1 (F), Albino	Timepoint 7 months	444	165, 5 (F), Brown	Timepoint 15 months
207	231, 2 (F), Albino	Timepoint 7 months	444	271, 2 (F), Brown	Timepoint 15 months
207	341, 1 (F), Brown	Timepoint 7 months	444	273, 4 (F), Brown	Timepoint 15 months
207	342, 2 (F), Brown	Timepoint 7 months	444	274, 5 (F), Brown	Timepoint 15 months
207	406, 6 (F), Brown	Timepoint 7 months	444	281, 4 (F), Albino	Timepoint 15 months
207	384, 7 (F), Brown	Timepoint 7 months	445	169, 3 (M), Brown	Timepoint 15 months
207	459, 10 (F), Brown	Timepoint 7 months	445	174, 3 (M), Albino	Timepoint 15 months
208	184, 1 (M), Brown	Timepoint 7 months	445	175, 1 (F), Black	Timepoint 15 months
208	185, 2 (M), Brown	Timepoint 7 months	445	176, 2 (F), Black	Timepoint 15 months
208	187, 4 (M), Brown	Timepoint 7 months	499	411, 3 (M), Brown	Removed: skinny, wound
208	419, 5 (F), Albino	Timepoint 7 months	560	420, 1 (F), Black	Natural death
308	507, 4 (M), Brown	Removed: testicular tumor	570	243, 3 (F), Brown	Removed: skinny
311	530, 5 (M), Brown	Natural death	570	220, 4 (M), Brown	Removed: abdominal tumor
312	532, 2 (M), Albino	Natural death	584	365, 2 (F), Albino	Removed: ovarian tumor
322	171, 5 (M), Brown	Removed: skinny, dehydrated	608	471, 8 (M), Brown	Removed: large spleen
352	192, 2 (F), Brown	Natural death	613	610, 3 (M), Brown	Natural death
401	288, 4 (F), Brown	Removed: enlarged spleen	620	539, 3 (M), Brown	Natural death
406	321, 3 (F), Brown	Natural death	626	376, 1 (F), Black	Natural death
442	246, 1 (M), Brown	Removed: enlarged spleen	626	383, 6 (F), Brown	Removed: tumor eye
431	564, 1 (M), Brown	Timepoint 15 months	626	462, 3 (F), Albino	Removed: skinny
435	524, 1 (M), Black	Timepoint 15 months	626	474, 3 (M), Albino	Removed: testicular tumor
435	525, 2 (M), Black	Timepoint 15 months	627	216, 2 (F), Albino	Removed: very skinny
444	527, 2 (M), Brown	Timepoint 15 months	627	252, 1 (F), Brown	Natural death
444	347, 1 (M), Brown	Timepoint 15 months	637	247, 2 (M), Brown	Removed: skinny
444	349, 3 (M), Brown	Timepoint 15 months	646	540, 4 (M), Brown	Removed: large spleen
444	505, 2 (M), Brown	Timepoint 15 months	647	538, 2 (M), Brown	Natural death

653	573, 2 (M), Brown	Timepoint 23 months	667	213, 1 (F), Brown	Timepoint 23 months
656	537, 1 (M), Brown	Timepoint 23 months	667	200, 2 (F), Brown	Removed: impaired activity
656	542, 6 (M), Brown	Timepoint 23 months	667	201, 3 (F), Brown	Removed: enlarged abdomen, large spleen
656	543, 7 (M), Brown	Timepoint 23 months	671	622, 4 (M), Brown	Timepoint 23 months
658	473, 2 (M), Albino	Removed	713	191, 1 (F), Brown	Removed: ovarian tumor
659	409, 1 (M), Brown	Removed: large spleen	751	181, 5 (F), Brown	Natural death
666	257, 2 (M), Brown	Timepoint 23 months	819	566, 3 (M), Brown	Natural death
666	259, 4 (M), Brown	Timepoint 23 months	860	591, 4 (M), Brown	Removed
666	261, 6 (M), Brown	Timepoint 23 months	888	319, 1 (F), Brown	Natural death
666	232, 3 (F), Albino	Timepoint 23 months	889	287, 3 (F), Brown	Natural death
666	239, 1 (F), Black	Timepoint 23 months	911	368, 1 (M), Brown	Removed: skinny, impaired activity
666	514, 2 (F), Black,	Timepoint 23 months	919	608, 1 (M), Brown	Natural death
666	522, 1 (F), Albino	Timepoint 23 months	923	385, 1 (F), Albino	Natural death
667	195, 1 (M), Brown	Timepoint 23 months	936	523, 2 (F), Albino	Natural death
667	248, 3 (M), Brown	Timepoint 23 months	966	285, 1 (F), Brown	Natural death
667	407, 1 (M), Black	Timepoint 23 months	1001	179, 3 (F), Brown	Natural death
667	200, 2 (F), Brown	Timepoint 23 months	1051	499, 1 (F), Brown	Natural death
667	201, 3 (F), Brown	Timepoint 23 months			

Individual data of each mice in the lifespan analysis; age (days), animal number, sex, coat color and indication of death. The ages of the mice were exactly 6.6-7.8±0.2, 14.5-15.6±0.1 and 21.8-22.5±0.2 months±SD. Lifespan analysis: n<sub>males</sub>=48, n<sub>females</sub>=51, including lifespan study: n<sub>males</sub>=20, n<sub>females</sub>=22 and behavioral timepoints: n<sub>M7</sub>=7, n<sub>M15</sub>=10, n<sub>M22</sub>=10, n<sub>F7</sub>=9, n<sub>F15</sub>=12, n<sub>F22</sub>=7.

**Supplementary Table 2. Additional physiological parameters.**

		<b>BL</b>	<b>Body Temp</b>	<b>BMD</b>	<b>BMC</b>	<b>Fat mass</b>	<b>Lean mass</b>
		<b>(cm)</b>	<b>(°C)</b>	<b>(mg/cm<sup>2</sup>)</b>	<b>(g)</b>	<b>(g)</b>	<b>(g)</b>
MALES	7	10.8 ± 0.11 <sup>2</sup>	36.7 ± 0.18 <sup>1,2</sup>	56.5 ± 1.34	0.62 ± 0.03	9.4 ± 1.2 <sup>2</sup>	25.5 ± 0.83 <sup>1,2</sup>
	15	10.7 ± 0.11	37.2 ± 0.11 <sup>1,2</sup>	57.1 ± 1.05	0.65 ± 0.02 <sup>2</sup>	9.9 ± 1.2 <sup>2</sup>	25.4 ± 0.98 <sup>1,2</sup>
	22	10.8 ± 0.11	36.5 ± 0.13 <sup>1,2</sup>	56.3 ± 1.13	0.65 ± 0.02	9.9 ± 1.7	28.7 ± 0.46 <sup>1,2</sup>
FEMALES	7	10.3 ± 0.09 <sup>2</sup>	38.4 ± 0.16 <sup>1,2</sup>	56.8 ± 0.84	0.56 ± 0.02	11.0 ± 0.8 <sup>2</sup>	20.0 ± 0.59 <sup>1,2</sup>
	15	10.5 ± 0.11	37.6 ± 0.10 <sup>1,2</sup>	49.3 ± 4.58	0.51 ± 0.04 <sup>2</sup>	14.1 ± 1.8 <sup>2</sup>	20.7 ± 0.85 <sup>1,2</sup>
	22	11.0 ± 0.12	37.7 ± 0.18 <sup>1,2</sup>	56.8 ± 1.35	0.63 ± 0.04	10.6 ± 1.6	23.5 ± 1.02 <sup>1,2</sup>

Body composition and core temperatures were analyzed by DEXA and rectal thermometer. <sup>1</sup>Statistical significance within sex, <sup>2</sup>statistical significance between sexes. Body length <sup>2</sup> $p_{M7-F7}=0.0040$ , body temperature <sup>1</sup> $p_{M7-15}=0.055$ ,  $p_{M7-15}=0.055$ ,  $p_{F7-15}=0.0017$ ,  $p_{M15-22}=0.00072$ ,  $p_{F15-22}=0.017$ ; <sup>2</sup> $p_{M7-F7}>0.0001$ ,  $p_{M15-F15}=0.0044$ ,  $p_{M22-F22}=0.00013$ , bone mineral density (BMD), bone mineral content (BMC) <sup>2</sup> $p_{M15-F15}=0.0060$ , fat mass <sup>2</sup> $p_{M7-F7}=0.0023$ ,  $p_{M15-F15}=0.0012$  (Welch's t-test), and lean mass <sup>2</sup> $p_{M7-F7}=0.0002$ ,  $p_{M15-F15}=0.0006$ ,  $p_{M22-F22}=0.0007$ , <sup>1</sup> $p_{M7-22}=0.0061$ ,  $p_{M15-22}=0.0095$ ,  $p_{F7-22}=0.011$ ,  $p_{F15-22}=0.023$  (Mann-Whitney).  $n_{M7}=7$   $n_{M15}=10$   $n_{M22}=10$   $n_{F7}=9$   $n_{F15}=12$   $n_{F22}=7$  Mean ± SEM.

**Supplementary Table 3. Order of behavioral tests at 7-, 15-, and 22-month-old timepoints.**

	MALES			FEMALES		
	7	15	22	7	15	22
FST	OFT	DEXA	OFT	DEXA	OFT	
OFT	DEXA	OFT	FST	FST	PAT	
PAT	PAT	FST	PAT	PAT	DEXA	
DEXA	FST	PAT	DEXA	OFT	FST	

Behavioral and physiological analyses: forced swim test (FST), activity box – open field test (OFT), shuttle-box passive avoidance test (PAT), dual energy X-ray absorptiometry (DEXA).  $n=7-12$  per group