Laboratory mouse housing conditions can be improved using common environmental enrichment without compromising data

Viola André[1]*, **Christine Gau**[2], **Angelika Scheideler**[1], **Juan A. Aguilar-Pimentel**[2], **Oana V. Amarie**[2,3], **Lore Becker**[2], **Lillian Garrett**[2,3], **Wolfgang Hans**[2], **Sabine M. Hölter**[2,3], **Dirk Janik**[2,4], **Kristin Moreth**[2], **Frauke Neff**[2,4], **Manuela Östereicher**[2], **Ildiko Racz**[2,5], **Birgit Rathkolb**[2,6,7], **Jan Rozman**[2,7], **Raffi Bekeredjian**[8], **Jochen Graw**[3], **Martin Klingenspor**[9,10], **Thomas Klopstock**[11,12,13], **Markus Ollert**[14,15], **Carsten Schmidt-Weber**[16], **Eckhard Wolf**[6], **Wolfgang Wurst**[3,12,13,17,18], **Valérie Gailus-Durner**[2], **Markus Brielmeier**[1], **Helmut Fuchs**[2], **Martin Hrabé de Angelis**[2,7,19]

1 Research Unit Comparative Medicine, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany, 2 German Mouse Clinic, Institute of Experimental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany, 3 Institute of Developmental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany, 4 Institute of Pathology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany, 5 Institute of Molecular Psychiatry, University of Bonn, Bonn, Germany, 6 Institute of Molecular Animal Breeding and Biotechnology, Gene Center, Ludwig-Maximilians-Universität München, Munich, Germany, 7 German Center for Diabetes Research (DZD), Neuherberg, Germany, 8 Department of Cardiology, University of Heidelberg, Heidelberg, Germany, 9 Molecular Nutritional Medicine, Else Kröner-Fresenius Center, Technische Universität München, Freising-Weihenstephan, Germany, 10 ZIEL—Center for Nutrition and Food Sciences, Technische Universität München, Freising, Germany, 11 Department of Neurology, Friedrich-Baur-Institute, Klinikum der Ludwig-Maximilians-Universität München, Munich, Germany, 12 Deutsches Institut für Neurodegenerative Erkrankungen (DZNE), Site Munich, Munich, Germany, 13 Munich Cluster for Systems Neurology (SyNergy), Adolf-Butenandt-Institut, Ludwig-Maximilians-Universität München, Munich, Germany, 14 Department of Dermatology and Allergy Center, Odense Research Center for Anaphylaxis (ORCA), Odense University hospital, University of Southern Denmark, Odense C, Denmark, 15 Department of Infection and Immunity, Luxembourg Institute of Health (LIH), Esch-sur-Alzette, Luxembourg, 16 Center of Allergy & Environment (ZAUM), Technische Universität München and Helmholtz Zentrum München, Munich, Germany and Member of the German Center for Lung Research (DZL), Gießen, Germany, 17 Technische Universität München, Freising-Weihenstephan, Chair of Developmental Genetics, c/o Helmholtz Zentrum München, Neuherberg, Germany, 18 Max Planck Institute of Psychiatry, Munich, Germany, 19 Chair of Experimental Genetics, School of Life Science Weihenstephan, Technische Universität München, Freising, Germany

* viola.andre@helmholtz-muenchen.de

Abstract

Animal welfare requires the adequate housing of animals to ensure health and well-being. The application of environmental enrichment is a way to improve the well-being of laboratory animals. However, it is important to know whether these enrichment items can be incorporated in experimental mouse husbandry without creating a divide between past and future experimental results. Previous small-scale studies have been inconsistent throughout the literature, and it is not yet completely understood whether and how enrichment might endanger comparability of results of scientific experiments. Here, we measured the effect on means and variability of 164 physiological parameters in 3 conditions: with nesting material with or without a shelter, comparing these 2 conditions to a "barren" regime without any enrichments. We studied a total of 360 mice from each of 2 mouse strains (C57BL/6NTac

---

and DBA/2NCrl) and both sexes for each of the 3 conditions. Our study indicates that enrichment affects the mean values of some of the 164 parameters with no consistent effects on variability. However, the influence of enrichment appears negligible compared to the effects of other influencing factors. Therefore, nesting material and shelters may be used to improve animal welfare without impairment of experimental outcome or loss of comparability to previous data collected under barren housing conditions.

Author summary

Adequate housing of laboratory animals is essential to guarantee their well-being. From a scientific perspective, physically and mentally healthy animals also contribute to increased validity and reproducibility of experimental results. The choice of nesting material or shelter type, referred to as environmental enrichment, may influence how laboratory animals perform species-specific behaviors. Consequently, changes in these nesting and shelter materials could influence scientific results by, for example, increasing variability in measured characteristics. Whether studies using different environmental enrichment materials can be compared is currently questioned. Our study shows that simple, species-specific environmental enrichment in the form of nesting material alone or in combination with a shelter did not consistently increase variability of physiological parameters in mice. Differences in parameter average values appeared to be of minor biological relevance when compared to the effects of other environmental factors. These simple environmental enrichment devices may therefore be applied to improve the housing environment of laboratory mice without compromising data validity or comparability.

Introduction

The provision of species-appropriate environmental enrichment—which can be defined as additions to the cage environment that allow natural motivated behaviors enabling the animals to control their environment [1]—is generally promoted as a way to improve animal welfare [1,2] and is also legally requested within the European Union by Directive 2010/63/EU [3]. However, there are various kinds of enrichment items available, and their imprudent application might interfere with the comparability of scientific results as even "seemingly minor alterations in the environment can have significant effects on experimental outcomes" [4]. A study by Macri et al. suggests that the adoption of environmental enrichment according to Directive 2010/63/EU might strongly influence the conclusions drawn from pharmacological and behavioral studies. In their study, they tested a synthetic cannabinoid compound and concluded that "whether the compound shall be considered a cannabinoid agonist may strongly depend on the specific conditions in which mice are reared" [5]. This underlines the potential risk following the implementation of inconsiderate enrichment strategies without preceding evaluation.

The term environmental enrichment is widely applied and also includes experimental paradigms, where intensive environmental enrichment strategies are used to explore effects of a more complex environment [6]. This so-called "super-enrichment"—as, for example, described in a protocol by Slater et al. [7]—induces behavioral [8], emotional [9,10], physiological [11], and neurobiological [10,12,13] changes in mice compared to barren housing. Moreover, such environmental enrichment can improve pathological conditions. For example,

enrichment can suppress tumor growth and reduce adiposity [7] and alleviate the intensities of various phenotypes in animal models (see [14] for an overview).

Besides such use of (super)enrichment as an experimental tool, simple enrichment is regularly used in laboratory animal facilities to ensure health and well-being of the animals and to meet their physiological and ethological needs as much as possible. It is thus applied as a refinement strategy according to the Three Rs (3Rs) of Russell and Burch [15]. The results of studies that determine the effects of those more commonly used forms of enrichment are varied.

Some authors found that nesting material influenced behavioral parameters of mice [16] or affected the scientific outcome in a well-described mouse model for allergic asthma [17]. Shelters as enrichment also altered motor coordination and some behavioral parameters in mice [18]. Furthermore, several studies revealed that even simple forms of enrichment, like nesting material [19], shelter [20], shelter combined with a scaffolding [21], or labyrinths [22], affected aggression and stress-related parameters in male mice of certain strains. Others found no effects of nesting material [23] or a shelter [24] on behavior and some physiological parameters.

Moreover, there has been much discussion about possible consequences of environmental enrichment on variability in data. Concern has been expressed by some authors that the response to and experience with environmental enrichment might differ between individuals so that environmental enrichment might lead to a higher variability in physiological parameters [25–28]. This might add to the individual variability within groups [29] and thereby increase the number of animals needed to reach statistical power. On the other hand, a multilaboratory study by Wolfer et al. found that environmental enrichment did not increase within-group variability in 19 out of 20 parameters of 4 behavioral tests in female mice [30]. Richter et al. describe a different approach and suggested that a conscious standardized heterogenization of environmental conditions might indeed increase within-experiment variation but then might also lead to lower between-experiment variation and therefore might improve external validity of experiments. An efficient heterogenization strategy is yet to be determined, though [31–33]. But while an environmental heterogenization strategy might give information about generalizability, according to van der Staay et al., there is the risk that subtle effects might be missed [34]. The choice of study design (strict standardization versus heterogenization) is therefore heavily dependent on the scientific questions a study aims to answer [34].

Taken together, despite a diverse body of literature, the assessment of the actual influence of enrichment remains difficult because the existing literature is, in large part, based on (single-cohort) studies that examine environmental enrichment effects under very specific conditions in respect to examined strain, sex, enrichment device, and parameters. But direction and size of effects seem to vary depending on the type and combination of enrichment [35], the mouse strain [36,37], sex [38], time and duration of enrichment [39], and which parameters were studied [26]; the applicability across rodent species, strains, sexes, and ages is uncertain [40]. To prevent unclear or spurious results and the need for higher numbers of animals due to increased variability within groups, a systematic evaluation of enrichment strategies is crucial [41,42], and environmental enrichment interventions should be "carefully selected, thoroughly defined, and purposefully used" [4]. This is of special importance in light of the ongoing debate about reproducibility because certain environmental factors might account for irreproducible rodent experiments [43,44] and variation or lack of standardization in environmental enrichment strategies might contribute to problems with reproducibility in preclinical research [4]. Kent et al. conclude in the ACLAM position statement on reproducibility that "it is incumbent on laboratory animal veterinarians and the scientific community to define elements of study design that affect experimental reproducibility" [45].

Screens	Age (weeks) / methods	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Screens	methods															
Behaviour	Open field			■												
	Acoustic startle response & PPI				■											
Neurology	Modified SHIRPA, grip strength			■												
	Rotarod				■												
Nociception	Hot plate						■									
Dysmorphology	Anatomical observation						■									
Energy metabolism	Indirect calorimetry, NMR							■								
Clinical chemistry	IpGTT									■						
Cardiovascular	awake ECG, Echocardiography										■					
Eye	Scheimpflug imaging, virtual drum										■					
Neurology	ABR (auditory brain stem response)												■			
Dysmorphology	DEXA												■			
Clinical chemistry	Clinical chemical analysis, hematology															■
Allergy	BIOPLEX ELISA (Ig concentration)															■
Pathology	Macro- and microscopic analysis															■

Fig 1. Workflow for the screens of this study. DEXA, dual-energy x-ray absorptiometry; IpGTT, intraperitoneal glucose tolerance test; NMR, nuclear magnetic resonance; PPI, pre-pulse inhibition; SHIRPA, Smithkline Beecham, MRC Harwell, Imperial College, the Royal London Hospital Phenotype Assessment.

https://doi.org/10.1371/journal.pbio.2005019.g001

Our study approach was therefore to investigate the effect of 2 simple, commonly used, and easily applicable enrichment items—namely, a shelter and nesting material (nestlet) versus nestlet alone or none of the above—on means and variability (expressed by coefficients of variation [CVs]) of a wide range of physiological parameters in a systematic, highly standardized study design (see Fig 1). To this end, we wanted to find out whether those environmental enrichments would change results of future and follow-up studies compared to the former standard of barren housing. For our purposes, we also specifically wished to ascertain whether environmental enrichment alters the measuring system of the German Mouse Clinic [46–48]. We examined male and female wild-type mice of 2 commonly used inbred mouse strains, C57BL/6NTac (B6) and DBA/2NCrl (D2). The study was conducted in 3 replicates using 3 independent cohorts of mice to give information about variability.

Results

The aim of this study was to evaluate the influence of simple environmental enrichment on the mean and variability of a broad range of physiological parameters in 2 strains of mice. The parameters we chose covered the areas of behavior, dysmorphology, neurology, clinical chemistry and hematology, eye, allergy, energy metabolism, pain perception, cardiovascular health, and pathology. Details of the parameters can be depicted from Fig 2 (quantitative parameters) and S1 Table, S2 Table, and S1 Text (qualitative parameters). To gain information about variability, 3 independent cohorts of mice were measured.

Environmental enrichment effects on mean values of quantitative physiological traits

We analyzed the influence of environmental enrichment on the mean of our parameters using linear models with "main factors": enrichment, sex, and cohort. We used a variable selection
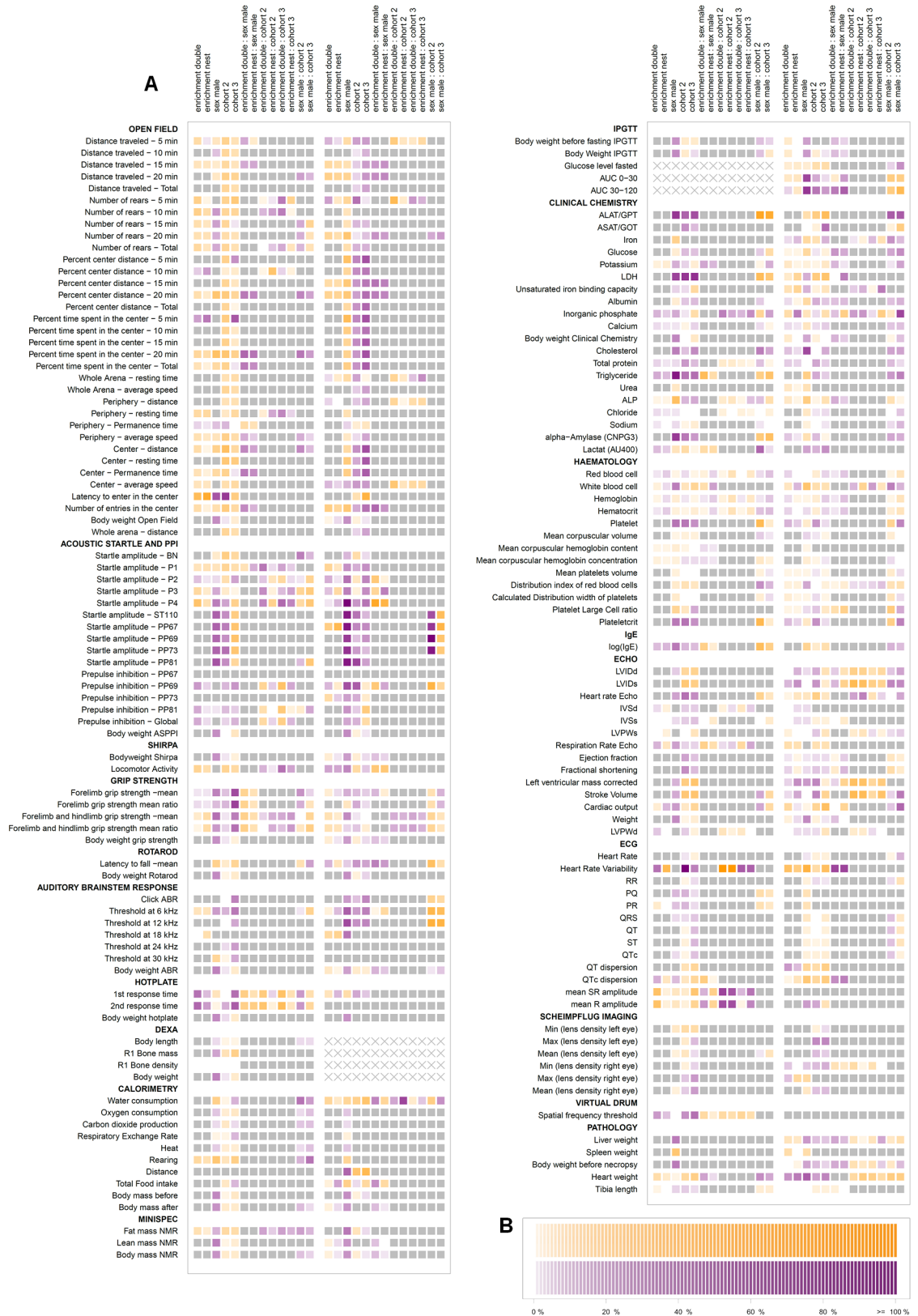
Fig 2. Influence of main factors (enrichment, sex, cohort) on mean of parameters in 2 mouse strains. (A) The heatmap shows the results for all metric parameters (rows) and influencing factors and their double interactions (columns). The B6 and D2 data are

on the left and right of each column, respectively. The influence of main factors was evaluated using linear models. A variable selection using the BIC was performed. Reference categories are enrichment none, sex female, and the first cohort. The differences between means of parameters are expressed as color-coded estimators β of the respective model in percent of the intercept of the model. The estimator β describes the change induced by the respective factors. The intercept is the mean of the parameter of the reference group and is set as 100%. Grey shade: variable selection using the BIC yielded nonrelevant influencing factor, thus no effect can be assumed; violet shade: factor is relevant for the model, i.e., mean is increased compared to reference group; yellow shade: factor is relevant for the model, i.e., mean is decreased compared to reference group; strength of color reflects size of difference. "X" marks parameters that were not analyzed. (B) Color grading scheme: the color grading illustrates the change in percent of the intercept; the strength of color reflects the size of difference. ABR, auditory brain stem response; ALAT/GPT, Alanine aminotransferase/Glutamat pyruvat transaminase; ALP, Alkaline phosphatase; ASAT/GOT, Aspartate aminotransferase/Glutamat oxalacetat transaminase; ASR, acoustic startle response; AU400, name of clinical chemistry analyzer; AUC, area under the curve; BIC, Bayesian Information Criterion; BN, background noise; CNPG3, substrate of test reaction (2-Chloro-4-Nitrophenyl-α-D-Maltotriosid); DEXA, dual-energy x-ray absorptiometry; HR, heart rate; HRV, heart rate variability; IgE, immunoglobulin E; IpGTT, intraperitoneal glucose tolerance test; LDH, Lactat-dehydrogenase; LVIDd, left ventricular end-diastolic internal diameter; LVIDs, left ventricular end-systolic internal diameter; LVPW, left ventricular posterior wall; P1-P4, startle amplitude following pulse of 67 (P1), 69 (P2), 73 (P3), 81 (P4) dB; PP (67, 69, 73, 81), startle amplitude and PPI following startle pulse (110 dB) with preceding pre-pulse of 67, 69, 73, 81 dB; PPI, pre-pulse inhibition; qNMR, quantitative nuclear magnetic resonance; QTc, corrected QT; R1, region analyzed (whole body excluding the skull); SHIRPA, Smithkline Beecham, MRC Harwell, Imperial College, the Royal London Hospital Phenotype Assessment; ST 110, acoustic startle response at 110 dB.

https://doi.org/10.1371/journal.pbio.2005019.g002

using the Bayesian Information Criterion (BIC) to determine differences between the experimental groups. The reference group was the unenriched females from cohort 1.

The influence of main factors (enrichment, sex, cohort) on the mean of all metric parameters and their double interactions is shown as a heatmap (Fig 2) with strength of color reflecting the size of difference. If the variable selection yielded a variable as nonrelevant, it was not included in the model, and the parameter was marked as grey (i.e., no difference) in the heatmap (Fig 2).

To better illustrate the effects of enrichment compared to the effects of factors sex and cohort, we computed smoothed histograms (Fig 3) of parameters that were influenced by the respective main factors.

For an exemplary evaluation of the biological relevance of our findings in the open field test, we prepared boxplots of raw data for the 4 main parameters on which we normally focus. They are "distance traveled total" (index of locomotor activity), "number of rears total" (index of exploratory activity), "percent center distance total," and "center permanence time" (indices of anxiety-related behavior). Furthermore, in each plot we included data of more than 200 B6 wild-type animals, which were measured as control mice in phenotyping projects of the German Mouse Clinic (GMC). Because these B6 were the same age and used in the same timespan as the mice of this study, the data can serve as a biological range for B6 mice (Fig 4).

Raw data in form of boxplots (S1 Fig) and individual values (S1 Data) for every parameter are provided in Supporting information.

Some aspects of our analysis are illustrated below by considering the 2 different mouse strains used (B6 and D2) individually.

Specific analysis of the means of the physiological parameters for the B6 strain

In B6 mice, the means of 69 out of 161 parameters were changed by enrichment; a higher number of parameters was affected by sex (118/161) and cohort (152/161).

Overall, percentage differences of the mean of the respective groups and their control were distinctly higher for factor sex and cohort than for factor enrichment (Fig 3A). Considerably fewer parameters were influenced by enrichment than by the other main factors as can be seen by the smaller area under the curve for factor enrichment compared to factors sex and cohort. Both enrichment curves were narrower than the curves of the other main factors. This

a

b
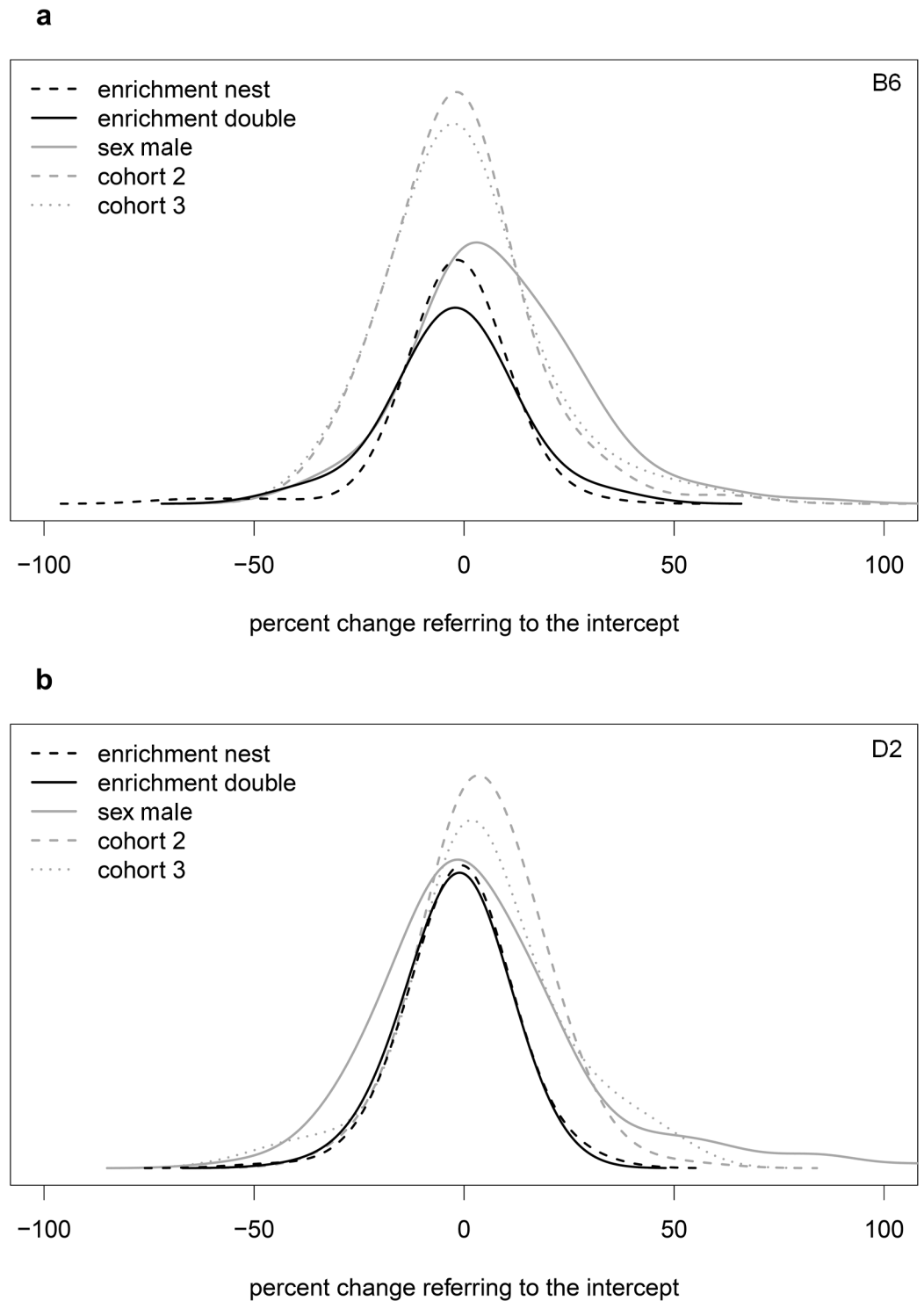
Fig 3. Smoothed histogram showing the influence of main factors (enrichment, sex, cohort) on means in 2 strains. The influence is expressed as differences in percent of intercept for B6 (a) and D2 (b) mice. The intercept is the mean of the parameter of the reference group. The area under the curve is proportional to number of parameters, for which the variable selection using the BIC yielded relevance (yellow and violet shading in Fig 2); nonselected parameters (grey in the heatmap of Fig 2) are not included (number of selected factors for B6 [a]: enrichment nest: 69; enrichment double: 69; sex male: 118; cohort 2: 152; cohort 3: 152; number of selected factors for D2 [b]: enrichment nest: 88; enrichment double: 88; sex male: 144; cohort 2: 135; cohort 3: 135). The height of the histogram is calculated as density multiplied by total number of parameters and is proportional to number of parameters that fall in the respective bin of the x-axis.

https://doi.org/10.1371/journal.pbio.2005019.g003

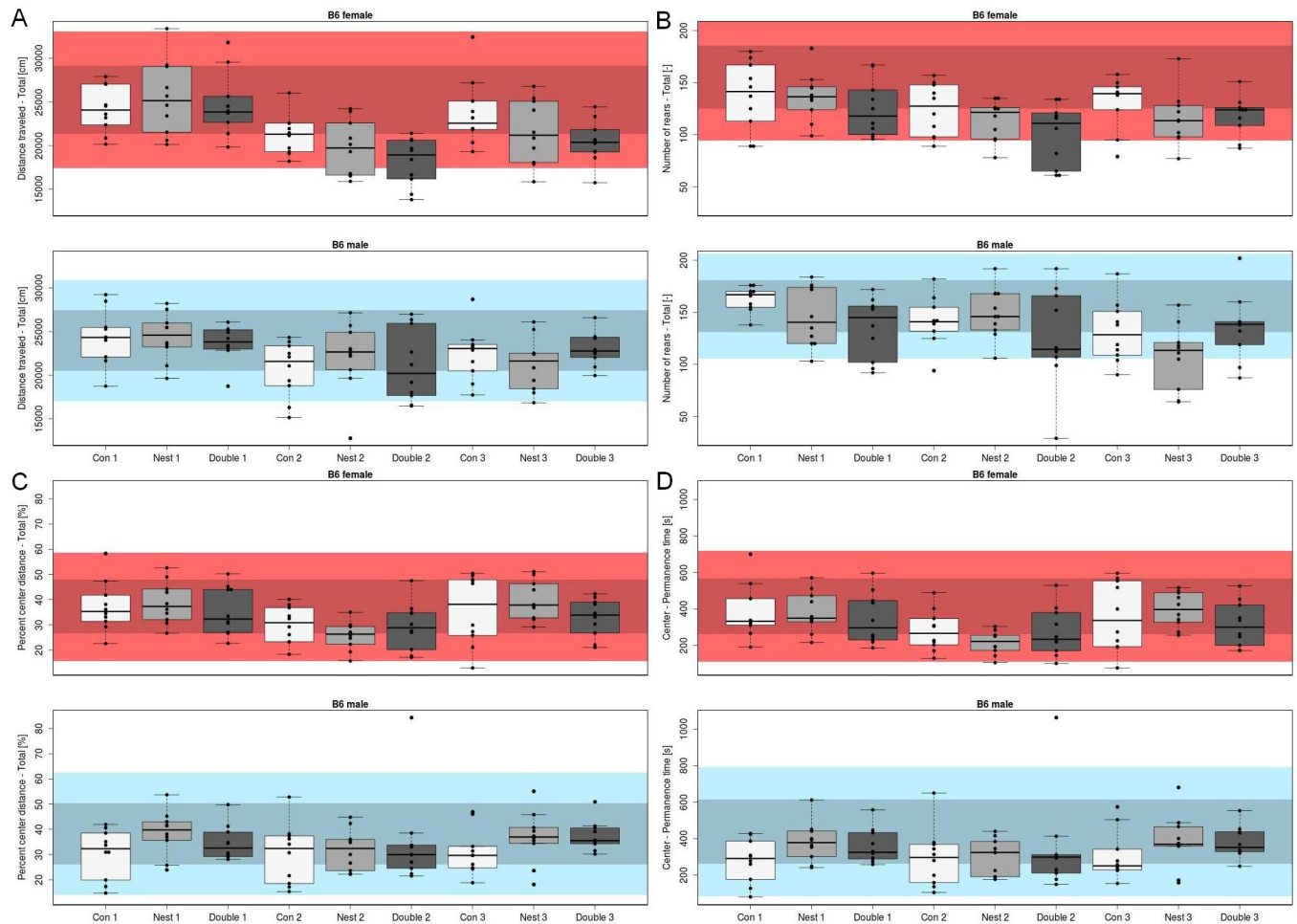**Fig 4. Raw data of selected parameters on the background of the biological range for B6.** Raw data of the 4 main parameters of the open field test: (A) "distance traveled total," (B) "number of rears total," (C) "percent center distance total," and (D) "center permanence time" as box- and whisker-plots. The box represents 25th percentile, median, and 75th percentile; the length of whiskers is maximally the 1.5-fold interquartile range but is determined by the last value within this range. All individual values are shown for each experimental group ("con," nest, double) and every cohort (1, 2, 3) for female and male mice in the upper and lower plot, respectively, for each selected parameter (A–D). The range in the background gives 1 SD (dark shading) and 2 SD (bright shading) of >200 reference B6 female (red) and male (blue) mice. The reference mice were same-aged wild-type control mice from other phenotyping projects of the GMC and were measured within the same timespan as the mice used in this project. "con", control; GMC, German Mouse Clinic.

https://doi.org/10.1371/journal.pbio.2005019.g004

indicates that the influence of enrichment was considerably smaller than that of the other main factors.

A closer look at single research areas in the heatmap in Fig 2 reveals that some were rather robust against housing effects. The measured parameters of the following tests were mainly unchanged by enrichment compared to controls in B6: rotarod (neurology), Scheimpflug imaging (eye), auditory brain stem response (ABR; neurology), indirect calorimetry, and quantitative nuclear magnetic resonance (qNMR; energy metabolism). Other tests were more susceptible to effects of enrichment; in the open field test (behavior), 20 out of 34 parameters were influenced by nest and double enrichment. The percentage difference between mean of nest- and double-enriched groups and mean of controls ranged between 1% and 15% in 17 out of 20 (double enrichment) and 18 out of 20 (nest enrichment) parameters. Two of the 4 main parameters of the open field test ("number of rears total" and "center permanence time") were affected by enrichment. For those parameters, most values lay within the biological range of

B6 ([Fig 4]). As expected, there were sex differences for many parameters. Although our tests were performed under strictly standardized conditions, there were differences between the cohort replicates for a wide range of 152 out of 161 parameters.

Specific analysis of the means of the physiological parameters for the D2 strain

For D2 mice, a difference between enriched and nonenriched mice could be found in 88 out of 160 parameters. For the other main factors, 144 out of 160 (sex) and 135 out of 160 (cohort) parameters were affected by the respective factors.

Overall, as in B6, fewer parameters were influenced by factor enrichment than by other main factors, which is expressed by the smaller area under the curve for factor enrichment in comparison to the other main factors ([Fig 3B]). Curves of double and nest enrichment were narrower than those of other main factors, which indicate that the influence of enrichment, on average, was smaller than the influence of factors cohort and sex.

Again, [Fig 2] reveals that some tests appeared to be less sensitive towards effects of environmental enrichment than others, for example, Scheimpflug imaging and virtual drum (eye), indirect calorimetry (energy metabolism), immunoglobulin E (IgE; allergy), and ECG (cardiovascular). As in B6, factor sex also influenced a large part of parameters with partially prominent effects in D2 mice. Factor cohort also influenced many parameters regarding mean values; 135 out of 160 parameters were changed by cohort in D2 mice.

Environmental enrichment effects on variability of measured parameters (expressed as CVs)

To evaluate the impact of environmental enrichment on variability, we used bootstrapped samples of the original data. Therefore, we drew a sample from each subgroup 1,000 times. As a comparable measure for variability, the CV was used.

Bootstrapped CV values were then analyzed with linear models, using enrichment, sex, and cohort as main factors.

The influence of the main factors on bootstrapped CVs is shown as a heatmap ([Fig 5]). Results of the bootstrap method are summarized by computed confidence intervals for estimators (β), and bootstrapped CVs were then classified into the following 3 categories: β includes 0, thus no effect is assumed (grey); confidence interval for β is greater than 0, i.e., bootstrapped CVs are increased compared to reference group (violet); confidence interval for β is below 0, i.e., bootstrapped CVs are decreased compared to reference group (yellow).

Specific analysis of the variability of the physiological parameters

In B6 mice, CVs of 84 out of 161 and 93 out of 161 parameters were affected (either increased or decreased) by double enrichment and nest enrichment, respectively, compared to controls. For D2 mice, double enrichment influenced CVs of 101 out of 160 parameters, and nest enrichment affected 91 out of 160 parameters in comparison to controls. Overall, no distinct patterns could be observed that hinted towards a general increase or decrease of CVs of parameters in a certain test due to factor enrichment. Rather, within tests, CVs of individual parameters were increased, decreased, or not changed concomitantly ([Fig 5]).

For the other main factors, CVs of 107 out of 161 (sex), 114 out of 161 (cohort 2), and 112 out of 161 (cohort 3) parameters were influenced in B6; in D2, CVs of 114 out of 160 (sex), 101 out of 160 (cohort 2), and 114 out of 160 (cohort 3) parameters were changed.

Fig 5. Influence of main factors (enrichment, sex, cohort) on bootstrapped CVs in 2 strains. The heatmap shows the results for all metric parameters (rows) and influencing factors (columns) for B6 (left column) and D2 (right column). Reference categories are enrichment none, sex female, and the first cohort. Results of the bootstrap method are summarized by computed confidence intervals for estimators (β), and bootstrapped CVs are then classified into the following 3 categories: confidence interval for β includes 0, thus no effect is assumed (grey); confidence interval for β is greater than 0, i.e., bootstrapped CVs are increased compared to reference group (violet); confidence interval for β is below 0, i.e., bootstrapped CVs are decreased compared to reference group (yellow). "X" mark parameters that were

not analyzed. ABR, auditory brain stem response; ALAT/GPT, Alanine aminotransferase/Glutamat pyruvat transaminase; ALP, Alkaline phosphatase; ASAT/GOT, Aspartate aminotransferase/Glutamat oxalacetat transaminase; ASR, acoustic startle response; AU400, name of clinical chemistry analyzer; AUC, area under the curve; BIC, Bayesian Information Criterion; BN, background noise; CNPG3, substrate of test reaction (2-Chloro-4-Nitrophenyl-α-D-Maltotriosid); CV, coefficient of variation; DEXA, dual-energy x-ray absorptiometry; HR, heart rate; HRV, heart rate variability; IgE, immunoglobulin E; IpGTT, intraperitoneal glucose tolerance test; LDH, Lactat-dehydrogenase; LVIDd, left ventricular end-diastolic internal diameter; LVIDs, left ventricular end-systolic internal diameter; LVPW, left ventricular posterior wall; P1-P4, startle amplitude following pulse of 67 (P1), 69 (P2), 73 (P3), 81 (P4) dB; PP (67, 69, 73, 81), startle amplitude and PPI following startle pulse (110 dB) with preceding pre-pulse of 67, 69, 73, 81 dB; PPI, pre-pulse inhibition; qNMR, quantitative nuclear magnetic resonance; QTc, corrected QT; R1, region analyzed (whole body excluding the skull); SHIRPA, Smithkline Beecham, MRC Harwell, Imperial College, the Royal London Hospital Phenotype Assessment; ST 110, acoustic startle response at 110 dB.

https://doi.org/10.1371/journal.pbio.2005019.g005

Overall, no clear indication could be found that factor enrichment induced higher CVs neither in B6 nor in D2. However, enrichment (double and nest) influenced variability in fewer parameters than sex and cohort in both strains of mice.

Categorical data

No effects of different housing conditions were found on categorical and qualitative data relevant for detection of abnormalities (in Smithkline Beecham, MRC Harwell, Imperial College, the Royal London Hospital Phenotype Assessment [SHIRPA] of the neurology screen, see S1 Table; in morphological examination of the dysmorphology screen, see S1 Text; in histopathological examination of the pathology screen, see S2 Table). Data of these tests were not included in the linear model analysis.

Discussion

The aim of this study was to investigate the effects of commonly used environmental enrichment on a comprehensive range of physiological parameters that cover key experimental procedures of medical research. Three independent cohorts of mice of 2 strains and both sexes with 3 different housing conditions were measured in a highly standardized study design.

Mean values of about half of our quantitative parameters were affected by enriched housing in both strains (B6: 43%; D2: 55%). However, the differences found were mostly small, and the biological relevance still has to be interpreted separately for each parameter. For example, in our study, two-thirds of the parameters of the open field test were changed by enrichment in B6 mice. But many of those parameters are correlated with each other, so that if one is changed, the others change concurrently, which might seem to inflate the number of affected parameters. Moreover, the differences of most parameters were rather small (1%–15%). To further evaluate the biological relevance of our findings in the open field test, we compared results of the 4 main parameters ("distance traveled total," "number of rears total," "percent center distance total," and "center permanence time"), on which we normally focus with the biological range of B6 mice. Two of the 4 mentioned parameters of the open field test ("number of rears total" and "center permanence time") were affected by enrichment. However, these effects could not be observed in all 3 cohorts, and even for these metrics, most of the measured values also lay within the biological range of B6 mice. This suggests that the effects of factor enrichment were within the regular variation that can be seen between different cohorts.

It has already been shown that large effects—e.g., strain differences in behavioral testing—could be reproduced in the environment of different laboratories despite differences in absolute values [49,50]. The fact that we did not find biologically relevant effects of simple

enrichment on, e.g., behavior, while others did [16] might indicate that the effects were too small to be found consistently.

On the other hand, we observed that some tests and parameters were robust towards the influence of environmental enrichment, as they were not changed. For some parameters, our results provide the same indication as studies in which no effects of enrichment were found in rotarod [16], body weight [27,51], food intake [51,52], or liver and spleen weight [27] in B6 mice. Other studies observed changes of body weight [23,52] and food intake [23] due to certain kinds of environmental enrichment, but—unlike them—our results did not hint towards changes in body weight. The differential findings between the studies can in part be explained by different study designs because factors like type and combination of enrichment [35], examined strain [36,37], sex [38], time and duration of enrichment [39], and studied parameters [26] influence the effects of environmental enrichment.

In our study, however, all those factors were standardized, and other environmental factors seemed to influence parameters equally or even more so than factor enrichment. These are represented by factor cohort, which changed 94% (B6) and 84% (D2) of examined parameters. Compared to factor enrichment, influence of cohort was apparent for more parameters, and on average, the observed effect was also stronger. Cohort effects have already been described as temporal variation [53] in phenotyping studies and can be attributed to differing body weights [54], seasonal variation [55], uncontrolled noise [56], sex of the experimenter [57], differing experimenters [58], and probably other factors that are yet unknown [4]. Tests in our study were conducted by differing experimenters of both sexes, which might have contributed to variation between cohorts [57,58]. Tail handling of mice has been shown to induce anxiety, which might have influenced behavioral results in our study [59,60], and battery testing itself can also induce additional noise [61] and add to the differences between cohorts. Influence of the above-discussed environmental factors attributing to cohort effects could also be accountable for differential study results that examine effects of simple enrichment in a single-cohort design. Furthermore, other (yet unknown) environmental factors might interfere with the measurements of the effects of enrichment. In a recent study, scientists failed to reproduce the finding that environmental enrichment decreased tumor growth [62] and also concluded that "other environmental factors are likely acting either in concert with or against environmental enrichment conditions to provide the variable results found" [63].

Altogether, even though we found that enrichment affects the mean of some parameters, its overall influence appeared to be of minor biological relevance in the background of the stronger environmental effects represented by the cohort. However, our study was designed to give a broad overview of possible effects of simple enrichment on a large number of physiological parameters of different research fields. Because we did not test an a priori hypothesis, our statistical analysis did not include a measure of statistical significance. Providing such a broad summary is helpful and can be used by other researchers to focus on single parameters of interest in confirmative studies. Furthermore, for logistical reasons, a blind outcome assessment was not possible in our study. However, most of the conducted tests are considered to be robust to subjective bias because animals are examined with the help of technical devices and parameters are digitally recorded and analyzed. We did not address the question, whether the used enrichment benefits animal welfare or provides a more realistic scenario in terms of resembling a natural environment. It is well known that barren housing conditions can cause impaired brain development and abnormal repetitive behaviors [13,64], which can compromise the validity of animal experiments and add variation [65]. However, the results we present here are, to our knowledge, the first systematic comparison of simple forms of enrichment with the former state-of-the-art—i.e., barren housing—on a large number of physiological parameters.

Apart from analyzing effects of common enrichment items on mean values, our second point of interest was in examining the effect on variability. We found that environmental enrichment influenced the CVs of 52% to 63% of parameters with no clear tendency towards an increase or a decrease. This is in concordance with other studies examining the effect of environmental enrichment on variability of data. Some studies found that variability of some parameters can be increased under enriched housing conditions [25,28] or rather decreased [58]. Others found no effects of housing on variability [11,66,67] or inconsistent results with effects on variability dependent on sex and studied parameters [26]. But the mentioned studies were single-cohort studies, which might account for inconsistency of results between studies. The only other multicohort study studied female mice only and found that within-group variability of several behavioral parameters was not affected by enriched housing [30].

Knowledge of possible influence of factors on variability is crucial to estimate whether reproducibility might be jeopardized. Concerns have been expressed that individual mice might interact differently with enrichment, which might lead to an increased variability of physiological parameters and therefore to higher animal numbers needed to obtain the appropriate statistical power in statistical evaluation [25,26]. If indeed enrichment led to higher variability, experimental results achieved with the same number of animals would not be reproducible after changing the housing conditions. This would yield an ethical conflict between reduction and refinement because enrichment is applied to enhance animal welfare (refinement), while higher animal numbers would represent a contrast to the principle of reduction. However, our study gave no clear indication that simple forms of environmental enrichment increase the variability of a broad range of physiological parameters. Moreover, it must be stressed that absolute replicability of results cannot be achieved because the environmental conditions cannot be fully reproduced despite standardization of environmental factors. However, relevant differences for hypothesis-driven comparisons should be reproducible over the small noise induced by differing environmental factors to be of external validity. As our study shows, simple environmental enrichment according to Directive 2010/63/EU adds only little to the noise between cohorts.

Broadly and comprehensively speaking, our data argue that simple environmental enrichment does not greatly vary relevant specific parameters of biological and medical enquiry. We conclude that nesting material and shelters can thus be liberally applied to improve laboratory animal welfare without skewing results, and new data from these conditions can be compared to past data that were collected in barren housing.

Material and methods

Ethics statement

All animal experiments were performed in compliance with German animal welfare law and were approved by the institutional animal care and use committee ("Committee for animal experiments and animal facility" of the Helmholtz Zentrum München) and by the District Government of Upper Bavaria (approval number: 55.2-1-54-2532-199-13).

Subjects and housing

A total number of 360 C57BL/6NTac (B6; Taconic, Denmark) and DBA/2NCrl (D2; Charles River, Germany) mice of both sexes were used in our study. Sixty mice (30 male, 30 female) of each strain were used in 1 set of examinations (see Table 1 for a full visual representation of the breakdown of the study design). This set of examinations was repeated twice so that 3 independent cohorts of 60 mice per strain were examined on the whole within a total timespan of 13 mo. Upon arrival, at 3 wk of age, animals were weighed, ear tagged, and split into 3 different groups (*n* = 10) following a stratified randomization scheme so that all groups had a similar

Table 1. Animal numbers used for this study.

Strain	B6						D2					
Experimental group	Control		Nest		Double		Control		Nest		Double	
Sex	M	F	M	F	M	F	M	F	M	F	M	F
Cohort 1	10	10	10	10	10	10	10	10	10	10	10	10
Cohort 2	10	10	10	10	10	10	10	10	10	10	10	10
Cohort 3	10	10	10	10	10	10	10	10	10	10	10	10

Altogether, 360 mice of 2 strains (C57BL/6NTac [B6], Taconic, Denmark; DBA/2NCrl [D2], Charles River, Germany) and both sexes were assigned to 3 experimental groups of different housing conditions (control, nest, double) according to the Table. One set of examinations is represented by each of the 3 cohorts (cohort 1, cohort 2, and cohort 3). Each cohort consisted of 12 groups with $n = 10$ animals per group. Due to technical reasons or death of animals during the course of the study, the numbers were reduced in some single groups or screens (see S2 Text). F, female; M, male.

https://doi.org/10.1371/journal.pbio.2005019.t001

body weight distribution at the beginning. All mice were housed in same-sex groups of 5, in type II polycarbonate cages in individually ventilated caging (IVC) systems (Tecniplast Green-line GM 500) with bedding (wood shavings, Altromin) and water and food ad libitum (standardized mouse diet, 1314, Altromin). The enrichment groups additionally had either a nestlet (PLEXX, Article ref. 14010)—which are cotton pads that mice can shred and use as nesting material (group "nest")—or a nestlet plus an orange plastic mouse igloo (PLEXX, Article ref. 13100) as a shelter (group "double"), whereas the "control" group had no enrichment items at all (pictures of the 3 housing conditions are shown in S1 Fig). Nestlets are commonly used in laboratory animal facilities and can additionally be applied to evaluate mouse welfare by nest complexity scoring as needed [68,69]. Cages were cleaned, and enrichment items were renewed weekly. On those occasions, mice were also weighed and examined to evaluate their health. The animal room had a controlled 12/12–h light/dark cycle (lights on at 6:00 AM), temperature (22 ± 2°C), and relative humidity (45%–65%).

General study design

At 3 wk of age, mice were imported into the animal facility and randomly assigned to the experimental groups ("control," "nest," "double"). From wk 9 to 21, mice were examined following the workflow of the primary screen for phenotypic analysis with minor adaptions (Fig 1).

Phenotype screens

The phenotyping screens were performed at the German Mouse Clinic, which offers a large-scale standardized and comprehensive phenotypic analysis of mice. In this study, mice were examined in the fields of behavior, dysmorphology, neurology, clinical chemistry and hematology, eye, allergy, energy metabolism, pain perception, cardiovascular health, and pathology (Fig 1). The phenotyping screens followed standardized examination protocols, as previously described (www.mouseclinic.de) [46–48]. In general, $n = 10$ mice per group (see Table 1) were examined, except for the eye screen ($n = 7$), cardiovascular screen ($n = 7$ in the first cohort of B6 mice, only for ECHO), and pathology screen ($n = 4$–5 for macroscopical analysis; $n = 1$–2 for histological analysis). Due to a few cases of unexpected death, animal numbers of groups were in some instances reduced to $n = 9$ and once to $n = 8$ (for details, see S2 Text).

Behavior

Open field. The test apparatus (ActiMot, TSE) was a transparent and infrared light-permeable acrylic test arena ($45.5 \times 45.5 \times 39.5$ cm internal measurements) surrounded by a

square-shaped frame with 2 pairs of light-beam strips. Lux levels were set at approximately 150 in the corners and 200 in the center of the arena. After at least 30 min of acclimatization to the test room, mice were carefully picked up by the base of the tail and were gently placed into the test arena to freely explore it for 20 min. After each trial, the arena was cleaned with a disinfectant (Pursept-A Xpress, Merz), which was used for the procedure in all cohorts. The following parameters were digitally measured: distance traveled; resting and permanence time; and speed of movement for the whole arena, the periphery, and the center as well as number of rearings. Rearing frequency, percentage distance traveled, and percentage time spent in the center, as well as the latency to first entry in the center and center entry frequency, were calculated.

Acoustic startle response and pre-pulse inhibition. Acoustic startle response (ASR) and pre-pulse inhibition (PPI) were assessed using a startle apparatus (Med Associates, St Albans City, VT). The protocol for PPI was based on the IMPReSS protocol from the International Mouse Phenotyping consortium (see https://www.mousephenotype.org/impress/protocol/176/7), adapted to the specifications of our startle equipment. Background noise was 65 dB, and startle pulses were bursts of white noise (40 msec). A session was initiated with a 5-min acclimation period followed by 5 presentations of leader startle pulses (110 dB) that were excluded from statistical analysis. Trial types for the PPI included 4 different pre-pulse intensities (67, 69, 73, 81 dB); each pre-pulse preceded the startle pulse (110 dB) by a 50-msec inter-stimulus interval. Each trial type was presented 10 times in random order, organized in 10 blocks, each trial type occurring once per block. Intertrial intervals varied from 20 to 30 s.

Neurology

Modified SHIRPA. Twenty-three parameters were examined to test for basic neurological function and general health by using a modification of the SHIRPA protocol [70]. The details of the procedure can be found elsewhere [71].

Grip strength. For measurement of muscle strength, a grip strength meter system was used (Bioseb, Chaville, France). Mice were lowered onto a metal grid that they were allowed to grasp with either 2 or 4 paws. They were then pulled away slowly, and the maximal force until they release the grid was measured by a force meter. Each mouse was measured 3 times within 1 min, and mean values were used as representative values for one individual mouse.

Rotarod. A rotarod (Bioseb, Chaville, France) was used to assess motor coordination, balance, and motor learning ability [72]. Mice were placed on the rotarod at an accelerating speed from 4 to 40 rpm for 5 min for 3 measurements, with 15 min between each trial, and latencies were recorded. Measurements were terminated when the mouse either fell off the rod, showed passive cycling, or after a maximum time of 5 min.

ABR. Measurement of the ABR was used to objectively evaluate hearing sensitivity of mice in anesthetized animals (i.p., 137 mg ketamine/6,6 mg xylazine/ kg body weight). Two electrodes were attached subcutaneously between mastoid and vertex to detect the physiological hearing curves produced by signal propagation in the brain stem following acoustic stimulation of different frequencies (6, 12, 18, 24, 30 Hz) in an ABR workstation (Tucker-Davis Technologies, Alachua, FL). Auditory threshold was determined as the minimum sound pressure level at the respective frequency that still triggered a typical ABR curve.

Nociception

Hot plate. The hot plate test was conducted as described recently [48]. In short, mice were placed on a metal plate that was maintained at $52 \pm 2 ^{\circ}$C, and type of the first 2 reactions and respective latencies were recorded for assessment of thermal sensitivity indicative for

nociceptive pathway integrity. Mice were taken off the plate as soon as they showed 2 signs or after maximum time of 30 s to avoid tissue lesion.

Dysmorphology

Anatomical observation.   The anatomical observation is a basic examination to screen for any morphological abnormalities and was performed as described previously [73].

DEXA.   DEXA analysis was executed during the same anesthesia as ABR following a standardized protocol [48]. For bone density measurement, a pDEXA Sabre X-ray Bone Densitometer (Norland Medical Systems, Basingstoke, Hampshire, UK) was used with the following settings: scan speed 20 mm/s, resolution 0.5 mm × 1.0 mm, histogram average width setting 0.020 g/cm^2.

For technical reasons, 1 cohort of D2 mice could not be measured with DEXA analysis and is therefore not included in linear model analysis. Data will be made accessible if requested.

Energy metabolism

Indirect calorimetry.   For assessment of energy metabolism, gas exchange (O_2 consumption, CO_2 production) was monitored for individual mice in an open-circuit respirometry system for 21 h (13:00 CET until 10:00 CET the next morning) with free access to food and water. Mice had no enrichment materials during this time. Respiratory exchange ratio could then be calculated. Additionally, locomotor activity and cumulative food intake were also recorded.

qNMR.   Body composition was measured with qNMR (MiniSpec LF50, Bruker Optics, Ettlingen, Germany). Body mass, lean tissue, and body fat were thereby assessed in live and awake mice. Both indirect calorimetry as well as qNMR were conducted as described earlier [48].

Clinical chemistry, hematology and allergy assessment

Intraperitoneal glucose tolerance test.   Intraperitoneal glucose tolerance test (IpGTT) was performed after 16 to 18 h of fasting. After initial assessment of body weight, basal glucose levels (after fasting) were determined by cutting off a maximum of 1 mm of the tail tip to receive a drop of blood. Glucose levels were analyzed with an Accu-Chek Aviva glucose analyzer (Roche/Mannheim). Then, mice were injected intraperitoneally with 2 g of glucose per kg body weight using a 20% glucose solution, a 25-gauge needle, and a 1-ml syringe. At 15, 30, 60, and 120 min after glucose injection, additional blood samples (1 drop each) were collected and used to determine blood glucose levels as described above. Repeated bleeding was induced by removing the clot from the first incision and massaging the tail of the mouse. For the analysis, the areas under the curve above basal glucose level for the first 30 min (AUC 0–30) and for the remaining 90 min (AUC 30–120) were calculated as described previously [74]. For technical reasons, parameters glucose level fasted, AUC 0–30, and AUC 30–120 had to be excluded for B6 mice and so were not included in the linear model analysis.

Blood sampling.   Blood samples were taken from isoflurane-anesthetized mice by puncturing the retrobulbar sinus with nonheparinized glass capillaries (1.0 mm in diameter; Neolab; Munich, Germany). After collecting the samples, mice were euthanized during narcosis by cervical dislocation or with CO_2. Blood samples were divided into 2 portions. The major portion was collected in a heparinized tube (Li-heparin, KABE; Numbrecht, Germany; Art.No. 078028). Cells and plasma were separated by a centrifugation step (10 min, 5,000 × g; 8˚C, Biofuge Fresco, Heraeus; Hanau, Germany). Plasma was used for measurement of immunoglobulin concentration and clinical chemistry assessment. The smaller portion was collected in an EDTA-coated tube (KABE, Art.No 078035) for hematological investigations.

Clinical chemistry, hematology, allergy assessment. Detailed information for measurements of clinical chemistry, hematology, and allergy assessment were described recently [48,75].

Cardiovascular health

Echocardiography. Left ventricular function was evaluated on nonanesthetized conscious mice by transthoracic echocardiography using a Vevo 2100 Imaging System (Visual Sonics) with a 30 MHz probe. Left ventricular parasternal short- and long-axis views were obtained in B-mode imaging, and left ventricular parasternal short-axis views were obtained in M-mode imaging at the papillary-muscle level. The short-axis M-mode images were used to measure left ventricular end-diastolic internal diameter (LVIDd), left ventricular end-systolic internal diameter (LVIDs), diastolic and systolic septal wall thickness (IVS), and left ventricular diastolic and systolic posterior wall (LVPW) thickness in 3 consecutive beats according to the American Society of Echocardiography leading edge method [76]. Some additional parameters such as fractional shortening, corrected left ventricular mass, stroke volume, and heart and respiratory rate were calculated from the above measured parameters [77].

For the first cohort of B6 mice, only n = 7 animals were examined.

Electrocardiography. ECGs were recorded in conscious mice with the ECGenie (Mouse Specifics, Boston, MA) and analyzed using e-Mouse software (Mouse Specifics, Boston, MA) The cardiac electrical activity was detected noninvasively through the animals' paws. The size and arrangement of the electrodes were configured to advance contact with 3 of the animals' paws to provide an ECG signal that is equivalent to Einthoven lead II. For each animal, intervals and amplitudes were evaluated from continuous recordings of at least 15 ECG signals. e-MOUSE software uses peak detection to calculate the heart rate (HR). HR variability (HRV) was calculated as the mean of the differences between sequential HRs. The software plots its interpretation of P, Q, R, S, and T for each beat so that unfiltered noise or motion artifacts are rejected. This is followed by calculations of the mean of the ECG time intervals for each set of waveforms. The corrected QT interval (QTc) was calculated by dividing the QT interval by the square root of the preceding RR interval. QT dispersion was measured as interlead variability of QT intervals. The QTc dispersion was calculated as the rate corrected–QT dispersion.

In the group female, double, B6 of the first cohort, 3 animals had to be excluded due to technical reasons.

Eye

Scheimpflug imaging. Images of corneas and lenses were taken with a Pentacam digital camera system (Oculus GmbH, Wetzlar, Germany). Pupils were widened by a drop of 0.5% atropine. Then, mice were held on a platform such that the vertical light slit (light source: LEDs, 475 nm) was orientated in the middle of the eyeball. Distance between eye and camera was fine-adjusted with the help of the provided software in order to guarantee optimal focus. Subsequently, the measurements were started manually. Mean density across the lens was quantified with the provided densitometry tool (Oculus GmbH, Wetzlar, Germany). Additionally, we conducted a qualitative examination of lens and cornea.

Virtual drum test. Vision tests were performed with the OptoMotry virtual optokinetic drum system (Cerebral Mechanics, Lethbridge, Canada) as described previously [78]. Briefly, a cylinder composed of a sine wave grating—drawn in 3D coordinate space on 4 computer monitors facing to form a square—rotates around the animal, which is placed on a platform in the middle of the drum. Visually unimpaired mice track the grating with reflexive head and neck movements (head tracking). Vision threshold of the tested mice was quantified by a

randomized simple staircase test. Rotation speed and contrast were set to 12.0 d/s and 100%, respectively.

For the eye screen, only n = 7 animals were examined.

Pathology

Pathological examination. Following blood sampling and euthanasia, mice underwent macroscopical pathological examination (http://eulep.pdn.cam.ac.uk/Necropsy_of_the_Mouse/index.php). After complete dissection, all organs were fixed in 4% buffered formalin and embedded in paraffin for histological examination. Four-μm–thick sections from skin, heart, muscle, lung, cerebrum, cerebellum, thymus, spleen, cervical lymph nodes, trachea, thyroid and parathyroid glands, adrenal glands, esophagus, stomach, intestine, liver, gall bladder, pancreas, kidneys, reproductive organs, and urinary bladder were cut and stained with hematoxylin–eosin (HE).

For macroscopical pathological examination, only n = 4–5 animals were examined; for histological examination, n = 1–2.

Statistical analysis

Our study aimed at answering the following questions: firstly, whether the mean of each parameter is influenced by enrichment and secondly, whether enrichment has an influence on the variability of the measurements. We considered sex and cohort as additional predictors for the possible alteration of mean and variability of the measurements.

The influence of enrichment on the mean values of the measured parameters was evaluated with linear models defined by

$$y = \beta_{Intercept} + \beta_{enrichment\ nest} + \beta_{enrichment\ double} + \beta_{sex\ male} + \beta_{cohort\ 2} + \beta_{cohort\ 3}$$
$$+ \ all\ double\ interactions + \varepsilon$$

A variable selection using the BIC was performed. The heatmap (Fig 2) shows the color-coded estimators (β) of the respective model in percent of the intercept of the model to provide a value which is comparable between the parameters. Nonselected influencing factors are marked in grey.

To evaluate the impact of enrichment on the variability, bootstrapped [79] samples of the original data were used. Therefore, 1,000 samples were drawn of size n_j from each group j (enrichment × sex × cohort combination). As a comparable measure for variability, the CV was used.

Linear models using the bootstrapped CV as outcome, and enrichment, sex, and cohort as influencing factors, were fitted. Because the interactions effects did not appear often as relevant in the linear models for the analysis of the mean values, only main effects have been included in the models for CV.

To summarize the results of the bootstrap method, empirical 95% confidence intervals for the estimators of all influencing variables of the linear models were computed. The results are summarized in 3 categories and are displayed in Fig 5, as follows: confidence interval includes 0 (grey; thus no effect is assumed), confidence interval is greater than 0 (violet), or confidence interval is below 0 (yellow).

All analyses were performed separately for each measured parameter and mouse strain. The fact that some of the parameters are highly correlated needs to be taken into consideration when interpreting the results.

Measures for significance were omitted due to the fact that the main purpose of the study was to get an overview about direction and magnitude of possible enrichment effects on parameters of different research fields. Providing such a broad summary is helpful and can be used by other researchers to focus on single parameters of interest in confirmative studies.

The statistic software R (version 3.0.2, R Foundation for Statistical Computing, Vienna, Austria) was used for all analyses and graphs. Dummy coding was used for all categorical factors in the linear models. Reference categories were as follows: enrichment: none; sex: female; and the cohort: 1.

Supporting information

S1 Data. Raw data of measured parameters (all screens).
(ZIP)

S1 Fig. Raw data of measured parameters (all screens). Raw data are shown as whisker-plots with the box representing 25th percentile, median, and 75th percentile. The length of whiskers is maximally the 1.5-fold interquartile range but is determined by the last value within this range. All individual values and each subgroup are shown. All individual values are shown for each experimental group (con, nest, double) and every cohort (1, 2, 3) for female and male mice in the upper and lower plot, respectively, for each of the 164 quantitative parameters.
(PDF)

S2 Fig. Housing conditions of the 3 experimental groups. (A) "Control" group without any items; (B) group "nest" with 1 cotton nestlet (PLEXX, Article ref. 14010); (C) group "double" with 1 cotton nestlet plus 1 plastic mouse igloo (PLEXX, Article ref. 13100).
(PNG)

S1 Table. Results of SHIRPA (neurology screen). Results are presented as number of animals receiving the parameter-specific score in combined cohorts and are described as follows: (x/y/z) with x = number of animals with score "0"; y = number of animals with score "1"; z = number of animals with score "2." Specifications for scores: body position: 0 = inactive; 1 = active; 2 = excessively active; tremor: 0 = absent; 1 = present; defecation: 0 = present; 1 = absent; transfer arousal: 0 = prolonged freeze; 1 = brief freeze; 2 = immediate movement; gait: 0 = fluid; 1 = abnormal; tail elevation: 0 = dragging; 1 = horizontal; 2 = elevated; startle response: 0 = no reaction; 1 = Preyer reflex; 2 = jumping; touch escape: 0 = no response; 1 = response to touch; 2 = flees prior to touch; trunk curl: 0 = absent; 1 = present; limb grasping: 0 = absent; 1 = present; pinna reflex: 0 = present; 1 = absent; urination: 0 = present; 1 = absent; contact righting reflex: 0 = present; 1 = absent; evidence of biting: 0 = no aggression; 1 = aggressive; vocalization: 0 = no; 1 = yes. SHIRPA, Smithkline Beecham, MRC Harwell, Imperial College, the Royal London hospital Phenotype Assessment.
(PDF)

S2 Table. Summary of background lesions in histopathological examination (pathology screen). Affected animals are shown as portion of total examined animals of combined cohorts.
(PDF)

S1 Text. Results of morphological examination (dysmorphology screen).
(PDF)

S2 Text. Reduction of animal numbers in single groups or screens.
(PDF)

Acknowledgments

Author Contributions

Conceptualization: Viola André, Angelika Scheideler, Markus Brielmeier, Helmut Fuchs, Martin Hrabé de Angelis.

Data curation: Viola André, Christine Gau, Manuela Östereicher.

Formal analysis: Christine Gau, Manuela Östereicher.

Funding acquisition: Martin Hrabé de Angelis.

Investigation: Juan A. Aguilar-Pimentel, Oana V. Amarie, Lore Becker, Lillian Garrett, Wolfgang Hans, Sabine M. Hölter, Dirk Janik, Kristin Moreth, Frauke Neff, Ildiko Racz, Birgit Rathkolb, Jan Rozman.

Methodology: Christine Gau, Manuela Östereicher.

Supervision: Angelika Scheideler, Raffi Bekeredjian, Jochen Graw, Martin Klingenspor, Thomas Klopstock, Markus Ollert, Carsten Schmidt-Weber, Eckhard Wolf, Wolfgang Wurst, Valérie Gailus-Durner, Markus Brielmeier, Helmut Fuchs.

Validation: Viola André, Christine Gau, Angelika Scheideler, Manuela Östereicher.

Writing – original draft: Viola André.

Writing – review & editing: Viola André, Christine Gau, Angelika Scheideler, Juan A. Aguilar-Pimentel, Oana V. Amarie, Lore Becker, Lillian Garrett, Wolfgang Hans, Sabine M. Hölter, Dirk Janik, Kristin Moreth, Frauke Neff, Manuela Östereicher, Ildiko Racz, Birgit Rathkolb, Jan Rozman, Raffi Bekeredjian, Jochen Graw, Martin Klingenspor, Thomas Klopstock, Markus Ollert, Carsten Schmidt-Weber, Eckhard Wolf, Wolfgang Wurst, Valérie Gailus-Durner, Markus Brielmeier, Helmut Fuchs, Martin Hrabé de Angelis.

References

1. Olsson IA, Dahlborn K. Improving housing conditions for laboratory mice: a review of "environmental enrichment". Lab Anim. 2002; 36(3):243–70. https://doi.org/10.1258/002367702320162379 PMID: 12144738.

2. Bayne K, Wurbel H. The impact of environmental enrichment on the outcome variability and scientific validity of laboratory animal studies. Revue scientifique et technique (International Office of Epizootics). 2014; 33(1):273–80. PMID: 25000800.

3. Directive 2010/63/EU of the european parliament and of the council of 22 September 2010 on the protection of animals used for scientific purposes, (2010).

4. Toth LA. The influence of the cage environment on rodent physiology and behavior: Implications for reproducibility of pre-clinical rodent research. Experimental neurology. 2015; 270:72–7. https://doi.org/10.1016/j.expneurol.2015.04.010 PMID: 25913548.

5. Macri S, Ceci C, Altabella L, Canese R, Laviola G. The Directive 2010/63/EU on animal experimentation may skew the conclusions of pharmacological and behavioural studies. Scientific reports. 2013; 3:2380. https://doi.org/10.1038/srep02380 PMID: 23924859.

6. Cao L, During MJ. What Is the Brain-Cancer Connection? Annual Review of Neuroscience. 2012; 35 (1):331–45. https://doi.org/10.1146/annurev-neuro-062111-150546 PMID: 22462541.

7. Slater AM, Cao L. A Protocol for Housing Mice in an Enriched Environment. Journal of visualized experiments: JoVE. 2015;(100):e52874. https://doi.org/10.3791/52874 PMID: 26131694.

8. Mesa-Gresa P, Perez-Martinez A, Redolat R. Environmental enrichment improves novel object recognition and enhances agonistic behavior in male mice. Aggress Behav. 2013; 39(4):269–79. https://doi.org/10.1002/ab.21481 PMID: 23588702.

9. Benaroya-Milshtein N, Hollander N, Apter A, Kukulansky T, Raz N, Wilf A, et al. Environmental enrichment in mice decreases anxiety, attenuates stress responses and enhances natural killer cell activity. Eur J Neurosci. 2004; 20(5):1341–7. https://doi.org/10.1111/j.1460-9568.2004.03587.x PMID: 15341605.

10. Zhu SW, Yee BK, Nyffeler M, Winblad B, Feldon J, Mohammed AH. Influence of differential housing on emotional behaviour and neurotrophin levels in mice. Behav Brain Res. 2006; 169(1):10–20. https://doi.org/10.1016/j.bbr.2005.11.024 PMID: 16406106.

11. Van de Weerd HA, Aarsen EL, Mulder A, Kruitwagen CL, Hendriksen CF, Baumans V. Effects of environmental enrichment for mice: variation in experimental results. J Appl Anim Welf Sci. 2002; 5(2):87–109. https://doi.org/10.1207/S15327604JAWS0502_01 PMID: 12738579.

12. Ehninger D, Kempermann G. Regional effects of wheel running and environmental enrichment on cell genesis and microglia proliferation in the adult murine neocortex. Cerebral cortex (New York, NY: 1991). 2003; 13(8):845–51. PMID: 12853371.

13. van Praag H, Kempermann G, Gage FH. Neural consequences of environmental enrichment. Nature reviews Neuroscience. 2000; 1(3):191–8. https://doi.org/10.1038/35044558 PMID: 11257907.

14. Nithianantharajah J, Hannan AJ. Enriched environments, experience-dependent plasticity and disorders of the nervous system. Nature reviews Neuroscience. 2006; 7(9):697–709. https://doi.org/10.1038/nrn1970 PMID: 16924259.

15. Russel WMS, Burch RL. The Principles of Humane Experimental Technique. London: Methuen; 1959.

16. Kulesskaya N, Rauvala H, Voikar V. Evaluation of social and physical enrichment in modulation of behavioural phenotype in C57BL/6J female mice. PLoS ONE. 2011; 6(9):e24755. https://doi.org/10.1371/journal.pone.0024755 PMID: 21931844.

17. Pasalic I, Bosnjak B, Tkalcevic VI, Jaran DS, Javorscak Z, Markovic D, et al. Cage enrichment with paper tissue, but not plastic tunnels, increases variability in mouse model of asthma. Lab Anim. 2011; 45(2):121–3. doi: la.2010.010109 [pii]. https://doi.org/10.1258/la.2010.010109 PMID: 21159849.

18. Coke-Murphy C, Buendia M, Saborido T, Stanwood G. Simple shelter-style environmental enrichment alters behavior in mice. TranslatNeurosci. 2014; 5(3):185–96. https://doi.org/10.2478/s13380-014-0228-4

19. Kaliste EK, Mering SM, Huuskonen HK. Environmental modification and agonistic behavior in NIH/S male mice: nesting material enhances fighting but shelters prevent it. Comp Med. 2006; 56(3):202–8. PMID: 16774129.

20. Van Loo P, Kruitwagen C, Koolhaas J, Van de Weerd H, Van Zutphen L, Baumans V. Influence of cage enrichment on aggressive behaviour and physiological parameters in male mice. Appl Anim Behav Sci. 2002; 76(1):65–81.

21. Marashi V, Barnekow A, Ossendorf E, Sachser N. Effects of different forms of environmental enrichment on behavioral, endocrinological, and immunological parameters in male mice. Horm Behav. 2003; 43(2):281–92. doi: S0018506X03000023 [pii]. PMID: 12694638.

22. Haemisch A, Voss T, Gartner K. Effects of environmental enrichment on aggressive behavior, dominance hierarchies, and endocrine states in male DBA/2J mice. Physiology & behavior. 1994; 56 (5):1041–8. PMID: 7824569.

23. Van de Weerd HA, Van Loo PL, Van Zutphen LF, Koolhaas JM, Baumans V. Nesting material as environmental enrichment has no adverse effects on behavior and physiology of laboratory mice. Physiology & behavior. 1997; 62(5):1019–28. PMID: 9333195.

24. Wirz A, Mandillo S, D'Amato FR, Giuliani A, Riviello MC. Response, use and habituation to a mouse house in C57BL/6J and BALB/c mice. Experimental animals / Japanese Association for Laboratory Animal Science. 2015; 64(3):281–93. https://doi.org/10.1538/expanim.14-0104 PMID: 25854626.

25. Eskola S, Lauhikari M, Voipio HM, Laitinen M, Nevalainen T. Environmental enrichment may alter the number of rats needed to achieve statistical significance. Scandinavian Journal of Laboratory Animal Science. 1999; 26(3):134–44.

26. Tsai PP, Stelzer HD, Hedrich HJ, Hackbarth H. Are the effects of different enrichment designs on the physiology and behaviour of DBA/2 mice consistent? Lab Anim. 2003; 37(4):314–27. https://doi.org/10.1258/002367703322389889 PMID: 14599306.

27. Tsai PP, Pachowsky U, Stelzer HD, Hackbarth H. Impact of environmental enrichment in mice. 1: effect of housing conditions on body weight, organ weights and haematology in different strains. Lab Anim. 2002; 36(4):411–9. https://doi.org/10.1258/002367702320389071 PMID: 12396284.

28. Mering S, Kaliste-Korhonen E, Nevalainen T. Estimates of appropriate number of rats: interaction with housing environment. Lab Anim. 2001; 35(1):80–90. https://doi.org/10.1258/0023677011911408 PMID: 11201291.

29. Loos M, Koopmans B, Aarts E, Maroteaux G, van der Sluis S, Verhage M, et al. Within-strain variation in behavior differs consistently between common inbred strains of mice. Mammalian genome: official

journal of the International Mammalian Genome Society. 2015; 26(7–8):348–54. https://doi.org/10.1007/s00335-015-9578-7 PMID: 26123533.

30. Wolfer DP, Litvin O, Morf S, Nitsch RM, Lipp HP, Wurbel H. Laboratory animal welfare: cage enrichment and mouse behaviour. Nature. 2004; 432(7019):821–2. https://doi.org/10.1038/432821a PMID: 15602544.

31. Richter SH, Garner JP, Wurbel H. Environmental standardization: cure or cause of poor reproducibility in animal experiments? Nat Methods. 2009; 6(4):257–61. https://doi.org/10.1038/nmeth.1312 PMID: 19333241.

32. Richter SH, Garner JP, Zipser B, Lewejohann L, Sachser N, Touma C, et al. Effect of population heterogenization on the reproducibility of mouse behavior: a multi-laboratory study. PLoS ONE. 2011; 6(1): e16461. https://doi.org/10.1371/journal.pone.0016461 PMID: 21305027.

33. Richter SH, Garner JP, Auer C, Kunert J, Wurbel H. Systematic variation improves reproducibility of animal experiments. Nat Meth. 2010; 7(3):167–8. doi: http://www.nature.com/nmeth/journal/v7/n3/suppinfo/nmeth0310-167_S1.html.

34. van der Staay FJ, Arndt SS, Nordquist RE. The standardization-generalization dilemma: a way out. Genes, brain, and behavior. 2010; 9(8):849–55. https://doi.org/10.1111/j.1601-183X.2010.00628.x PMID: 20662940.

35. Swetter BJ, Karpiak CP, Cannon JT. Separating the effects of shelter from additional cage enhancements for group-housed BALB/cJ mice. Neurosci Lett. 2011; 495(3):205–9. https://doi.org/10.1016/j.neulet.2011.03.067 PMID: 21457758.

36. van de Weerd HA, Baumans V, Koolhaas JM, van Zutphen LF. Strain specific behavioural response to environmental enrichment in the mouse. J Exp Anim Sci. 1994; 36(4–5):117–27. PMID: 7948063.

37. Abramov U, Puussaar T, Raud S, Kurrikoff K, Vasar E. Behavioural differences between C57BL/6 and 129S6/SvEv strains are reinforced by environmental enrichment. Neurosci Lett. 2008; 443(3):223–7. https://doi.org/10.1016/j.neulet.2008.07.075 PMID: 18687379.

38. Lin EJ, Choi E, Liu X, Martin A, During MJ. Environmental enrichment exerts sex-specific effects on emotionality in C57BL/6J mice. Behav Brain Res. 2010; 216(1):349–57. doi: S0166-4328(10)00578-4 [pii]. https://doi.org/10.1016/j.bbr.2010.08.019 PMID: 20732356.

39. Amaral OB, Vargas RS, Hansel G, Izquierdo I, Souza DO. Duration of environmental enrichment influences the magnitude and persistence of its behavioral effects on mice. Physiology & behavior. 2008; 93 (1–2):388–94. https://doi.org/10.1016/j.physbeh.2007.09.009 PMID: 17949760.

40. Toth LA, Kregel K, Leon L, Musch TI. Environmental enrichment of laboratory rodents: the answer depends on the question. Comp Med. 2011; 61(4):314–21. PMID: 22330246.

41. Baumans V, Clausing P, Hubrecht R, Reber A, Vitale A, Wyffels E, et al. FELASA Working Group Standardization of Enrichment. Working Group Report: FELASA; 2006. [Cited 7 March 2013]. Available from: http://www.felasa.eu/recommendations/reports/Standardization-of-enrichment/.

42. Benefiel AC, Dong WK, Greenough WT. Mandatory "enriched" housing of laboratory animals: the need for evidence-based evaluation. ILAR J. 2005; 46(2):95–105. PMID: 15775019.

43. Reardon S. A mouse's house may ruin experiments. Nature. 2016; 530(7590):264. https://doi.org/10.1038/nature.2016.19335 PMID: 26887470.

44. Troublesome variability in mouse studies. Nature neuroscience. 2009; 12(9):1075. https://doi.org/10.1038/nn0909-1075 PMID: 19710643.

45. Kent Lloyd KC, Niemi SM, Beaver BV, Berridge BR, Chamberlain PL, Clarke CL, et al. ACLAM position statement on reproducibility. Journal of the American Association for Laboratory Animal Science. 2016; 55(6):824–5. PMID: 27931326

46. Gailus-Durner V, Fuchs H, Becker L, Bolle I, Brielmeier M, Calzada-Wack J, et al. Introducing the German Mouse Clinic: open access platform for standardized phenotyping. Nat Methods. 2005; 2(6):403–4. https://doi.org/10.1038/nmeth0605-403 PMID: 15908916.

47. Fuchs H, Gailus-Durner V, Neschen S, Adler T, Afonso LC, Aguilar-Pimentel JA, et al. Innovations in phenotyping of mouse models in the German Mouse Clinic. Mammalian genome: official journal of the International Mammalian Genome Society. 2012; 23(9–10):611–22. https://doi.org/10.1007/s00335-012-9415-1 PMID: 22926221.

48. Fuchs H, Gailus-Durner V, Adler T, Aguilar-Pimentel JA, Becker L, Calzada-Wack J, et al. Mouse phenotyping. Methods. 2011; 53(2):120–35. https://doi.org/10.1016/j.ymeth.2010.08.006 PMID: 20708688.

49. Wahlsten D, Metten P, Phillips TJ, Boehm SL 2nd, Burkhart-Kasch S, Dorow J, et al. Different data from different labs: lessons from studies of gene-environment interaction. Journal of neurobiology. 2003; 54 (1):283–311. https://doi.org/10.1002/neu.10173 PMID: 12486710.

50. Mandillo S, Tucci V, Holter SM, Meziane H, Banchaabouchi MA, Kallnik M, et al. Reliability, robustness, and reproducibility in mouse behavioral phenotyping: a cross-laboratory study. Physiol Genomics.

2008; 34(3):243–55. doi: 90207.2008 [pii] https://doi.org/10.1152/physiolgenomics.90207.2008 PMID: 18505770.

51. Wirz A, Mandillo S, D'Amato FR, Giuliani A, Riviello MC. Response, use and habituation to a mouse house in C57BL/6J and BALB/c mice. Experimental animals / Japanese Association for Laboratory Animal Science. 2015. https://doi.org/10.1538/expanim.14-0104 PMID: 25854626.

52. Gurfein BT, Stamm AW, Bacchetti P, Dallman MF, Nadkarni NA, Milush JM, et al. The calm mouse: an animal model of stress reduction. Mol Med. 2012; 18:606–17. doi: molmed.2012.00053 [pii]. https://doi.org/10.2119/molmed.2012.00053 PMID: 22398685.

53. Karp NA, Speak AO, White JK, Adams DJ, Hrabe de Angelis M, Herault Y, et al. Impact of temporal variation on design and analysis of mouse knockout phenotyping studies. PLoS ONE. 2014; 9(10): e111239. https://doi.org/10.1371/journal.pone.0111239 PMID: 25343444.

54. Karp NA, Melvin D, Mott RF. Robust and sensitive analysis of mouse knockout phenotypes. PLoS ONE. 2012; 7(12):e52410. https://doi.org/10.1371/journal.pone.0052410 PMID: 23300663.

55. Delahunty KM, Horton LG, Coombs HF 3rd, Shultz KL, Svenson KL, Marion MA, et al. Gender- and compartment-specific bone loss in C57BL/6J mice: correlation to season? Journal of clinical densitometry: the official journal of the International Society for Clinical Densitometry. 2009; 12(1):89–94. https://doi.org/10.1016/j.jocd.2008.10.008 PMID: 19195621.

56. Turner JG, Parrish JL, Hughes LF, Toth LA, Caspary DM. Hearing in laboratory animals: strain differences and nonauditory effects of noise. Comp Med. 2005; 55(1):12–23. PMID: 15766204.

57. Sorge RE, Martin LJ, Isbester KA, Sotocinal SG, Rosen S, Tuttle AH, et al. Olfactory exposure to males, including men, causes stress and related analgesia in rodents. Nat Methods. 2014; 11(6):629–32. https://doi.org/10.1038/nmeth.2935 PMID: 24776635.

58. Lewejohann L, Reinhard C, Schrewe A, Brandewiede J, Haemisch A, Gortz N, et al. Environmental bias? Effects of housing conditions, laboratory environment and experimenter on behavioral tests. Genes, brain, and behavior. 2006; 5(1):64–72. https://doi.org/10.1111/j.1601-183X.2005.00140.x PMID: 16436190.

59. Hurst JL, West RS. Taming anxiety in laboratory mice. Nat Methods. 2010; 7(10):825–6. https://doi.org/10.1038/nmeth.1500 PMID: 20835246.

60. Gouveia K, Hurst JL. Optimising reliability of mouse performance in behavioural testing: the major role of non-aversive handling. Scientific reports. 2017; 7:44999. https://doi.org/10.1038/srep44999 PMID: 28322308.

61. Lad HV, Liu L, Paya-Cano JL, Parsons MJ, Kember R, Fernandes C, et al. Behavioural battery testing: evaluation and behavioural outcomes in 8 inbred mouse strains. Physiology & behavior. 2010; 99 (3):301–16. https://doi.org/10.1016/j.physbeh.2009.11.007 PMID: 19931548.

62. Cao L, Liu X, Lin EJ, Wang C, Choi EY, Riban V, et al. Environmental and genetic activation of a brain-adipocyte BDNF/leptin axis causes cancer remission and inhibition. Cell. 2010; 142(1):52–64. https://doi.org/10.1016/j.cell.2010.05.029 PMID: 20603014.

63. Westwood JA, Darcy PK, Kershaw MH. Environmental enrichment does not impact on tumor growth in mice. F1000Research [Internet]. 2013 Pmc3892918]; 2:[140 p.]. Available from: http://f1000researchdata.s3.amazonaws.com/manuscripts/1596/5df49036-9fb1-484c-823f-a199703c9bf7_1419%20-%20kershaw%20-%202-140.pdf.

64. Wurbel H. Ideal homes? Housing effects on rodent brain and behaviour. Trends in neurosciences. 2001; 24(4):207–11. PMID: 11250003.

65. Garner JP. Stereotypies and other abnormal repetitive behaviors: potential impact on validity, reliability, and replicability of scientific outcomes. ILAR J. 2005; 46(2):106–17. PMID: 15775020.

66. Augustsson H, van de Weerd HA, Kruitwagen CL, Baumans V. Effect of enrichment on variation and results in the light/dark test. Lab Anim. 2003; 37(4):328–40. https://doi.org/10.1258/002367703322389898 PMID: 14599307.

67. Baumans V, PLP VL, Pham TM. Standardisation of Environmental Enrichment for Laboratory Mice and Rats: Utilisation, Practicality and Variation in Experimental Results. Scand J Lab Anim Sci 2010. 2010; 37(2).

68. Jirkof P, Fleischmann T, Cesarovic N, Rettich A, Vogel J, Arras M. Assessment of postsurgical distress and pain in laboratory mice by nest complexity scoring. Lab Anim. 2013; 47(3):153–61. https://doi.org/10.1177/0023677213475603 PMID: 23563122.

69. Deacon RM. Assessing nest building in mice. Nature protocols. 2006; 1(3):1117–9. https://doi.org/10.1038/nprot.2006.170 PMID: 17406392.

70. Rogers DC, Fisher EM, Brown SD, Peters J, Hunter AJ, Martin JE. Behavioral and functional analysis of mouse phenotype: SHIRPA, a proposed protocol for comprehensive phenotype assessment.

Mammalian genome: official journal of the International Mammalian Genome Society. 1997; 8(10):711–3. PMID: 9321461.

71. Schneider I, Tirsch WS, Faus-Kessler T, Becker L, Kling E, Busse RL, et al. Systematic, standardized and comprehensive neurological phenotyping of inbred mice strains in the German Mouse Clinic. Journal of neuroscience methods. 2006; 157(1):82–90. https://doi.org/10.1016/j.jneumeth.2006.04.002 PMID: 16720049.

72. Jones BJ, Roberts DJ. A rotarod suitable for quantitative measurements of motor incoordination in naive mice. Naunyn-Schmiedebergs Archiv fur experimentelle Pathologie und Pharmakologie. 1968; 259(2):211. PMID: 4232694.

73. Fuchs H, Schughart K, Wolf E, Balling R, Hrabe de Angelis M. Screening for dysmorphological abnormalities—a powerful tool to isolate new mouse mutants. Mammalian genome: official journal of the International Mammalian Genome Society. 2000; 11(7):528–30. PMID: 10886017.

74. Rozman J, Rathkolb B, Neschen S, Fuchs H, Gailus-Durner V, Klingenspor M, et al. Glucose tolerance tests for systematic screening of glucose homeostasis in mice. Curr Protoc Mouse Biol. 2015; 5(1):65–84. https://doi.org/10.1002/9780470942390.mo140111 PMID: 25727201.

75. Neff F, Flores-Dominguez D, Ryan DP, Horsch M, Schroder S, Adler T, et al. Rapamycin extends murine lifespan but has limited effects on aging. The Journal of clinical investigation. 2013; 123 (8):3272–91. https://doi.org/10.1172/JCI67674 PMID: 23863708.

76. Sahn DJ, DeMaria A, Kisslo J, Weyman A. Recommendations regarding quantitation in M-mode echocardiography: results of a survey of echocardiographic measurements. Circulation. 1978; 58(6):1072–83. PMID: 709763.

77. Moreth K, Fischer R, Fuchs H, Gailus-Durner V, Wurst W, Katus HA, et al. High-throughput phenotypic assessment of cardiac physiology in four commonly used inbred mouse strains. Journal of comparative physiology B, Biochemical, systemic, and environmental physiology. 2014; 184(6):763–75. https://doi.org/10.1007/s00360-014-0830-3 PMID: 24788387.

78. Prusky GT, Alam NM, Beekman S, Douglas RM. Rapid quantification of adult and developing mouse spatial vision using a virtual optomotor system. Investigative ophthalmology & visual science. 2004; 45 (12):4611–6. https://doi.org/10.1167/iovs.04-0541 PMID: 15557474.

79. Efron B, Tibshirani RJ. An introduction to the bootstrap: CRC press; 1994.