Anna Goncerzewicz

Brain size and cognitive abilities of laboratory mice divergently selected for Basal or Peak Metabolic Rates

Ph.D. thesis
completed in the Laboratory of Emotions Neurobiology
of the Nencki Institute of Experimental Biology
Polish Academy of Sciences

SUPERVISORS:
dr hab. Ewelina Knapska, prof. of the Nencki Institute,
Nencki Institute of Experimental Biology, PAS
prof. dr hab. Marek Konarzewski, Institute of Biology,
University of Białystok

Warsaw, 2022

Acknowledgements

First and foremost I would like to express my sincere gratitude to prof. Marek Konarzewski – my PhD thesis supervisor, for giving me the opportunity to perform this project, for huge support and commitment as well as scientific supervision in the course of my research.

I would also like to express my deepest appreciation my promotor dr hab. Ewelina Knapska who accepted me as a member of Laboratory of Emotions' Neurobiology and created an opportunity to work in an inspiring scientific environment she created.

I would like to thank all past and present members of the Laboratory of Emotions' Neurobiology, especially Tomasz Górkiewicz, Ksenia Meyza, Tomasz Lebitko, Weronika Szadzińska for the encouragement and all support during my first steps conducting experiments, practical advice and for creating a friendly working atmosphere.

Additionally, special acknowledgements for Joanna Jędrzejewska-Szmek, Monika Malinowska and Agnieszka Kępczyńska for their cooperation and willingness to help in IC data analysis and all histological procedures.

Special thanks to my Family: Marek, Marcelina and Gabrysia for their unconditional approval, emotional support, motivation for further work and helping me to become a better version of myself every day.

This thesis goes out to my Mom.

I believe, You would be proud.

Table of content

1 Abbreviations

BMR	Basal Metabolic Rate
H-BMR	mice of the line type selected for high Basal Metabolic Rate
L-BMR	mice of the line type selected for low Basal Metabolic Rate
RB	random-bred line (non-selected mice)
PMR	mice of the line type selected for Peak Metabolic Rate (aka VO_{2max})
ET	the “expensive tissue” hypothesis
EB	the “expensive brain” hypothesis
LTP	long-term potentiation
fEPSPs	field excitatory postsynaptic potentials
IC	IntelliCage system
RFID	radio frequency based identification
CA	cognitive abilities
CA1	*Cornu Ammonis* subfield 1 of the hippocampus
CA2	*Cornu Ammonis* subfield 2 of the hippocampus
CA3	*Cornu Ammonis* subfield 3 of the hippocampus
DG	dentate gyrus subfield of the hippocampus
PFA	paraformaldehyde
PBS	phosphate buffer saline
PB	phosphate buffer
DiI	1,1′-dioctadecyl-3,3,3′,3′-tetramethylindocarbocyanine perchlorate dye
GFAP	glial fibrillary acidic protein
DAPI	4′,6-diamidino-2-phenylindole
DAB	3,3'-diaminobenzidine
BSA	bovine serum albumin
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
EDTA	ethylenediaminetetraacetic acid
RT	room temperature
o/n	overnight procedure
SEM	standard error of the mean
SE	standard error of covariance parameter estimation
p	statistical value of probability of significance
SAS	Statistical Analysis System

$\mathbf{s^2}$	covariance parameter estimate
$\boldsymbol{\chi^2}$	2 times the log likelihood from the model of a given dependent variable with only the fixed effects minus -2 times the log likelihood from the full model (with random factor)
N.E.	not estimable
n	number of animals included in the statistical analyses
F	generation of animals
$\boldsymbol{h^2}$	heritability
d	studied traits

All the gene and protein names are in accordance with guidelines of Hugo Gene Nomenclature Committee and UniProt database, respectively

2 Summary

The enlarged brains of homeotherms bring behavioral advantages but also incur high energy expenditures. Energy fueling evolutionary increase in brain size and enhanced cognitive abilities (CA) could come from two primary sources: according to the "expensive tissue" hypothesis postulated by (Aiello and Wheeler 1995), the evolution of a larger brain was made possible by a diet-related reduction in the size of the digestive tract and by increasing of quality (energy density) of the diet. Thereby, an evolutionary increase in brain size resulted from the brain-gut trade-off. The second hypothesis, dubbed the "expensive brain" hypothesis (Isler and van Schaik 2006), predicts that the energetic costs of an evolutionary increase in brain size were covered by increased total energy intake rather than energy savings on metabolically costly organs (such as the gut) or processes (reproduction or immunocompetence).

In my thesis, I asked a question: How were the energetic costs of an enlarged brain overcome in the course of evolution? To answer this question, I used the experimental evolution animal model consisting of the line types of Swiss Webster mice artificially selected for high (H-) or low (L-) Basal Metabolic Rate (BMR), maximal (VO_{2max}) metabolic rate (a.k.a. peak, PMR), and random bread lines (RB). The metabolism rates selected in the model are proxies of the traits implicated in the evolution of homeothermy. Thus, they are a prerequisite for the encephalization and exceptional CA of mammals, including humans. The H-BMR mice had bigger guts, but not brains, than mice of other line types. Yet, they were superior to the other line types in the cognitive tasks carried out in reward and avoidance learning contexts. Conversely, when subjected to the classical paradigm of contextual fear conditioning, the L-BMR mice lost fear response much faster than the mice of other line types (that is, their memory was inferior). Furthermore, the H-BMR mice had higher neuronal plasticity (indexed as the long-term potentiation, LTP). They also had increased numbers of neurons and dendritic spines in the hippocampus compared to their counterparts. Finally, the activity of cytochrome oxidase (CCO), a proxy of the number of neuronal mitochondria, was higher in the H-BMR mice than in other line types.

The results suggest that the evolutionary increase of CA in mammals was initially associated with increased BMR and brain plasticity, rather than a direct increase in brain size. Thus, an enlarged gut was not traded off for brain size. It could be that in the course of evolution, selection for increased total energy expenditures indirectly increased BMR and the metabolic rate of better connected and more plastic individual neurons, improving CA. Thus, my study

does not support the existence of the brain-gut trade-offs postulated by the ET hypothesis. Conversely, my results support the link between CA fueled by high brain metabolism reflected in H-BMR as proposed by the EB concept.

3 Streszczenie

Dużych rozmiarów mózg występujący u organizmów stałocieplnych niesie ze sobą ogromne korzyści związane ze zwiększeniem zdolności kognitywnych, ale wymaga zwiększonych nakładów energii. Może ona pochodzić z różnych źródeł. Według hipotezy „kosztownych tkanek" postulowanej przez (Aiello and Wheeler 1995), koszt utrzymania dużych rozmiarów mózgu jest kompensowany przez ograniczenie zużycia energii osiągane przez zredukowanie rozmiarów organów wewnętrznych takich jak jelita oraz poprzez przyjmowanego bardziej kalorycznego pokarmu. W związku z tym musiało dojść do ustanowienia kompromisu pomiędzy wielkością mózgu a wielkością jelit. Alternatywna hipoteza wyjaśniająca ewolucję rozmiaru mózgu nazwana hipotezą „kosztownego mózgu" (Isler and van Schaik 2006), która podważa powyższą hipotezę i zakłada, że zwiększone nakłady energii potrzebne do ewolucji mózgu i zdolności kognitywnych pochodzą ze zwiększenia puli energii dostępnej w środowisku (spożywanie większej ilości/bardziej kalorycznego pokarmu) oraz ze zmniejszenia zużycia energii przez procesy takie jak reprodukcja i immunokompetencja. Hipoteza „kosztownych tkanek" zakłada, że energia potrzebna na encefalizację pochodzi z jej relokacji, zmniejszenia jelit, a tym samym kosztów ich utrzymania względem bardziej kosztownego, większego mózgu. Natomiast hipoteza „kosztownego mózgu" wskazuje, że energia potrzebna do ewolucji zwiększonych rozmiarów mózgu nie wynikała z jej relokacji, a przez zwiększenie jej ilości w środowisku. Nie ma danych doświadczalnych, które pozwoliłyby na zweryfikowanie opisanych hipotez.

W związku z tym w swojej pracy stawiam pytanie: jak koszty związane z dużych rozmiarów mózgiem zostały przezwyciężone w toku ewolucji? Aby odpowiedzieć na postawione pytanie użyłam jako modelu myszy szczepu Swiss Webster pochodzących z eksperymentu selekcyjnego, który pozwolił na wyselekcjonowanie linii myszy o różnych poziomach metabolizmu, wysokiego (H-) oraz niskiego (L-) tempa metabolizmu podstawowego (BMR), maksymalnego (VO_{2max}) tempa metabolizmu wysiłkowego (PMR) oraz linii kontrolnych (RB). Cechy, które były podstawą selekcji są ściśle związane z ewolucją organizmów stałocieplnych i są wstępnym warunkiem encefalizacji oraz wyjątkowych zdolności kognitywnych ssaków, w tym ludzi.

W swoich badaniach wykazałam, że myszy wyselekcjonowane na wysokie tempo metabolizmu podstawowego (H-BMR) mają, w przeciwieństwie do innych badanych linii, większe organy wewnętrzne. U testowanych szczepów myszy nie zmienia się natomiast wielkość mózgu.

Myszy H-BMR przewyższały jednak myszy o niskim tempie metabolizmu (L-BMR) i zwierzęta kontrolne podczas testów apetytywnych oraz awersyjnych wykonywanych w systemie IntelliCage. Natomiast zwierzęta L-BMR słabiej niż inne linie myszy uczyły się podczas klasycznego warunkowania strachu. Badanie długotrwałego wzmocnienia synaptycznego (LTP) w ścieżce CA3-CA1 hipokampa wykazało najwyższą plastyczność synaptyczną u myszy H-BMR. Myszy te charakteryzowały się również większą liczbę neuronów i kolców dendrytycznych tworzących aktywne synapsy w hipokampie w stosunku do pozostałych badanych linii. Ponadto, myszy H-BRM miały podwyższoną aktywność oksydazy cytochromowej, co odzwierciedla zwiększoną aktywność mitochondriów.

Uzyskane wyniki wskazują, że ewolucyjne zwiększenie zdolności kognitywnych ssaków było związane ze zwiększeniem metabolizmu podstawowego oraz plastycznością neuronalnej, a nie poprzez wprost proporcjonalne zwiększenie wielkości mózgu. W związku z tym powiększenie organów wewnętrznych nie było ewolucyjnym kompromisem przy rozwoju większego mózgu. Niezbędna do tego procesu energia nie została uzyskana przez ograniczenie rozmiarów narządów wewnętrznych czy wzrostu.

Uzyskane rezultaty moich badań nie potwierdzają zatem hipotezy „kosztownych tkanek", wykazują natomiast związek pomiędzy zdolnościami kognitywnymi wykazanymi w testach behawioralnych a metabolizmem podstawowym zakładanymi przez hipotezę „kosztownego mózgu".

4 Introduction

4.1 Evolution of brain size from an energetic perspective

Increased brain size, relative to body mass, is the crucial characteristic of homeotherms (Armstrong and Bergeron 1985; Hofman 2015). High encephalization resulted in increased behavioral complexity but also incurred energy expenditures an order of magnitude higher than that of ectotherms (Pontzer et al. 2016). Furthermore, among homeotherms, relatively larger brains contribute disproportionally to energy expenditures. Indeed, brain metabolism accounts for ca. 8% of BMR in mice, but over 10% in primates, and up to an exceptional 20% of BMR in humans, even though in all those species brain constitutes only 2-3% of body mass (Erbslöh, Bernsmeier, and Hillesheim 1958; Mink, Blumenschine, and Adams 1981). An enduring question, then, is how the energetic costs of evolving a larger brain were overcome and what sort of energetic, anatomic, and physiological trade-offs and/or inherent positive associations were involved.

Any given trait or body function requires energy for its maintenance. Individual's ability to acquire energy is limited either by environmental resource limitations or physiological constraints, such as the digestive abilities of the gut (Thurber et al. 2019). Indeed, those limitations should bear a direct relationship to the brain size evolution because brain tissue is energetically expensive and need more energy to maintain than any other tissue in the body (Lewin 1982). Furthermore, brain energetic expenditures cannot be temporarily reduced in response to energy limitation (McKenna et al. 2012). The high proportion of energy that must be allocated to brain tissue may impose severe pressure on the evolution of brain size, despite the benefits of its increase. In the evolution of primates, those demands were met by higher caloric content of the diet (consisting of meat, fruits, and insects) or (in the case of humans) food that was easier to digest when processed by cooking (Wrangham and Conklin-Brittain 2003). However, the evolutionary increase in primates' brain size did not simply require higher caloric input. Since apes are incapable of torpor and/or hibernation (Biggar and Storey 2015), the energy supply to their brains must be continuous and without excessive periods of starvation. This requirement most likely posed a significant challenge to proto-humans evolving in the seasonal environments, where food availability is often too low to sustain energetic demands of large brains (van Woerden et al. 2012; Van Woerden, Van Schaik, and Isler 2010). The challenge of environmental seasonality was probably the main factor behind the evolution

of both the unique cognitive abilities of humans and their exceptional propensity to accumulate fat depots essential for tiding over food shortages (Zafon 2007). This peculiarity of human physiology is one of the fundamental evidences for the significance of energy expenditures in the evolution of the human brain.

4.2 Main hypotheses linking energetics and the evolution of brain size and function

Studies on brain size evolution identified several factors fostering encephalization. Among them, the most prominent ones are ecological (Parker and Gibson 1977; Milton 1981; Harvey, Clutton-Brock, and Mace 1980; Sol 2008), cognitive (Sol 2008; Allman, McLaughlin, and Hakeem 1993), and sociological factors (Pérez-Barbería, Shultz, and Dunbar 2007; Dunbar, n.d.; Whiten and Van Schaik 2007; Street et al. 2017; Jolly 1966), or foraging strategies (Morand-Ferron, Sol, and Lefebvre 2007; Burini and Leonard 2018; Eisenberg and Wilson 1978). However, since cognitive capabilities are commonly considered adaptive, why are they not ubiquitous? The answer to this question may lie in exceptionally high metabolic costs of maintenance of brain tissue. Two, partially complementary concepts have embraced this idea. The first one, known as the "expensive tissue" hypothesis (the ET hypothesis, thereafter), posits that increased encephalization was primarily possible thanks to funding metabolic costs of the enlarged brain by a reduction in energetically demanding gastrointestinal tract, which in turn was possible owing to increased cognitive abilities allowing for more efficient foraging (Aiello and Wheeler 1995). So, the ET hypothesis predicts a negative correlation between the brain and the gut size.

The ET hypothesis has been exceptionally well received as an explanation for the unique features of the evolution of the human brain. However, the ET scenario, in essence, refers to physiological mechanisms; therefore, it is difficult to test because such mechanisms are not directly preserved in a paleontological record. Its evolutionary significance can thus only be tested if they persist as general evolutionary principles also manifested in extant species. In particular, a strong assumption of the hypotheses explaining encephalization, including the ET hypothesis, is the existence of a positive association between enlarged brains and enhanced CA. Indeed, the ET hypothesis found support through studies on guppies (*Poecilia reticulata*) (Kotrschal et al. 2013), cichlid fishes (Tsuboi et al. 2018), frogs, and toads (Liao et al. 2016). Nevertheless, the ET hypothesis is hampered by many inconsistences and weaknesses. Several studies demonstrated a positive association between brain size and basal metabolic rate, BMR, e.g. (Bennett & Harvey, 1985; Isler & van Schaik, 2006; Mace et al., 1981). Since roughly 50%

of BMR reflects metabolic costs of maintenance of internal organs, including the gut (Konarzewski and Diamond 1995), a positive association between brain size and BMR is difficult to reconcile with the ET hypothesis. Furthermore, even though the ET concept was widely accepted among anthropologists, it has been questioned in a thorough comparative analysis of 100 mammalian species, including 23 primates (Navarrete, van Schaik, and Isler 2011). The ET hypothesis's major weakness is its overly narrow assumption of a direct relationship between brain size/mass and CA. Even more simplistic is the assumption that a 'bigger brain is better' without paying due attention to well-documented vast differences in neuronal density that can appear even within the same species (Carlo and Stevens 2013; Herculano-Houzel et al. 2014)

BMR, reflecting metabolic costs of gut maintenance, considered within the ET hypothesis framework, is not the only measure of metabolic expenditures likely to be linked with brain size. Researchers (Raichlen and Gordon 2011) recently found a positive association between exercise capacity (indexed as Peak Metabolic Rate, PMR) and brain size in a comparative analysis of non-human mammals. The existence of such a correlation was also supported by the results of Chappell et al.' (Chappell et al. 2007) study, which showed that brain size positively correlates with maximum voluntary aerobic capacity in gerbils. Furthermore, selection on maximum voluntary aerobic capacity in mice resulted in enlargement of the mid-brain and the dentate gyrus (Kolb et al. 2013). It also increased hippocampal neurogenesis (Rhodes et al. 2003), which may provide a mechanistic basis for the link between brain size and PMR. On the other hand, however, the same study failed to find a direct link between PMR and CA. Likewise, another study (Chrzaścik et al. 2014) did not detect such a link in the bank voles selected for high swim-induced aerobic metabolism, intensity of predatory behavior, and ability to grow on a low-quality herbivorous diet, which calls into question the association of the CA with thus selected traits.

The second hypothesis relating to the evolution of brain size and energetics is the "expensive brain" hypothesis (EB, thereafter) (Isler and van Schaik 2009). In contrast to the ET hypothesis, it does not predict the brain-gut trade-off and points to a positive association between metabolic rates and brain size. Thus, according to the EB scenario, the energetic costs of an evolutionary increase in brain size were fueled by increased total energy intake rather than energy savings achieved through evolutionary reduction of energy allocation to other metabolically expensive functions, such as body maintenance or production (growth and reproduction) (Isler and van Schaik 2009). The EB hypothesis was tested by Pontzer and collaborators (Pontzer et al. 2016),

who used doubly labeled water measurements of total energy expenditure (TEE; kcal day-1) in humans and apes. Pontzer et al. compared their BMR and Total Energy Expenditures (TEE) and concluded that human lineage had elevated metabolic rates, thereby providing more energy for larger brains and faster reproduction without reducing metabolic maintenance costs. Hominin brain evolution was therefore fueled in large part by increased mass-specific organ metabolic rates rather than anatomical or physiological trade-offs.

4.3 Energetic costs of neurophysiological processes

Glucose is an essential metabolic substrate and energy source in the mammalian brain (Mergenthaler et al. 2013). Oxygen metabolism provides energy essential for brain activities such as the maintenance of synaptic transmission and the resting potential of neurons (Attwell and Iadecola 2002; Shulman et al. 2004). The adenosine triphosphate (ATP) consumption by the major subcellular processes underlying signaling in the brain has been estimated for rat and human cerebral cortex (Attwell and Laughlin 2001). For example, fast excitatory synaptic transmission requires 1.4 × 105 ATP molecules per presynaptic stimulation to pump out an influx of Na^+ ions (Attwell and Laughlin 2001), which (assuming 1 Hz for such stimulation in the rat cortex) yields a metabolic rate of 8.4 × 106 ATP/min per spine (Karbowski 2019). Interestingly, recent data on cortical stimulation and energetics of synaptic transmission in rodents reveal that the small increase in the cortical metabolic rate is shared proportionally between neurons and astrocytes (Sonnay et al. 2018; 2015), suggesting that the metabolism of both neuronal and glial compartments are essential for synaptic function.

In larger mammalian species, a specific brain metabolism rate (metabolic rate per gram of tissue) parallels a decrease of neuronal density with increasing brain size (Herculano-Houzel 2011). With the smaller number of larger neurons, characterized by a reduced average number of abrupt changes in action potential (firing rate), larger brains need less energy per gram of tissue to maintain their function. However, it is unclear whether larger neurons are energetically more costly. Energy usage depends strongly on the action potential rate - an increase in activity of one action potential per cortical neuron increases oxygen consumption by 145 mL/100 g grey matter/h of oxygen consumption, as reported by Attwell & Laughlin (Attwell and Laughlin 2001). However, Herculano-Houzel (Herculano-Houzel 2011) showed that estimated glucose use per neuron is relatively constant, varying only by 40% across the six species of rodents and primates (including humans). Furthermore, average glucose usage per neuron does not correlate with neuronal density in any brain structure (Herculano-Houzel 2011). These observations

suggest that the whole brain's energy budget per neuron is fixed across species and brain sizes, which is difficult to reconcile with estimates of metabolic costs of neuronal activity (Attwell and Laughlin 2001; Harris, Jolivet, and Attwell 2012). In any case, however, remarkable energy consumption by the brain is roughly proportional to the number of neurons building a particular brain.

The brain's CA are determined not only by the number and/or size of neurons but also by the phenomenon of neuroplasticity. This term is used by neuroscientists, referring commonly to the brain's ability to change its structure and function in response to the changing patterns of incoming stimuli. This flexibility is incredibly important in learning and preserving new memories. The first concept linking higher mental abilities and neural activity was proposed by Polish scientist Jerzy Konorski in 1948 and almost simultaneously by Donald Hebb in 1949. They defined synaptic plasticity as a change in connectivity between neurons (Kandel 2012). At the single cell level, synaptic plasticity refers to synaptic connections changes. Neuroplasticity allows neurons to adjust their activities in response to new situations or changes in their environment. Changes in dendritic spine morphology and dynamics (the postsynaptic sites of most excitatory synapses in the brain) serve as good indicators of synaptic plasticity (Gray 1959; Yuste and Bonhoeffer 2001). Spines are specialized protrusions emerging from neuronal dendrites, with characteristic bulbous enlargements at their tips (spine heads). The spines are not static but actively move and alter their morphology continuously, even in the adult brain, reflecting the plastic nature of synaptic connections. Additionally, new spines could be formed during an animal's whole life, becoming functional synapses and eventually replacing the non-activated ones. They are firstly formed in early postnatal life, shaped up by the experience, and maintained into adulthood (Holtmaat and Svoboda 2009; Trachtenberg et al. 2002). Also, learning requires morphological changes in the neuronal connections and the formation of new synapses. Thus, another question about the portion of the total energy budget accounting for synaptic plasticity correlated with learning and memory appears. Based on neurophysiological and proteomic data for the rat brain (Karbowski 2019) it is estimated that, depending on the level of protein phosphorylation, the energy cost of synaptic plasticity constitutes energy used for fast excitatory synaptic transmission, typically adding 4.0 to 11.2%. Karbowski also discovered that longer memories require more energy to store, except for the parameters controlling the speed of molecular transitions (e.g., ATP-driven phosphorylation rate).

Neurotransmission is the process by which signaling molecules called neurotransmitters are released by the axon terminal of a presynaptic neuron and bind and react with the receptors on the dendrites of another postsynaptic neuron located a short distance. It is also an energy-demanding process. For instance, glutamate, the dominant excitatory neurotransmitter in the brain, released by ~90% of the neurons during excitation, diffuses across the synaptic cleft and is recognized by receptors on the postsynaptic neuron (Van Den Pol, Wuarin, and Dudek 1990; Magistretti 2009). Astrocytes take up glutamate via specific transporters that use the electrochemical gradient of Na^+ as a driving force, resulting in a tight coupling between glutamate and Na^+ uptake. Glutamine is subsequently released by astrocytes and taken up by neuronal terminals, where it is enzymatically reconverted to glutamate to replenish the neurotransmitter pool of glutamate. Both glutamine synthesis and the Na^+, K^+-ATPase require ATP (Magistretti and Allaman 2015; Magistretti et al. 1999). It has been shown that signal transmission is the most energy-consuming process in the brain. Most of the ATP is consumed during ion pumping, which maintains the resting potential. Moreover, the Na^+ ions that enter through the synaptic membrane to produce action potentials must be pumped out by the Na^+/K^+-ATPase. Since the costs of neural transmission are so high, one might expect that in the course of evolution they have been optimized with regard to both brain size and neuronal density (Harris, Jolivet, and Attwell 2012). To sum up, synaptic activity has tremendous implications for the energy costs of brain activity, and therefore, for the evolution of brain size and cognitive abilities.

4.4 Cognitive abilities vs. brain size

Almost all of the recent studies on cognitive abilities vs. brain size focused on aspects of complex behavior, under the assumption that performing of unusual or cognitively demanding tasks requires a larger brain (Kotrschal et al. 2013; Roth and Dicke 2005). Going further, we can ask the question whether there is a relationship between performance in behavioral tests and brain size and/or numbers of neurons in functionally relevant brain structures. ET and EB hypotheses conclude that larger brains translate to higher cognitive performance, better foraging behavior and adaptation to the environment. Up to now, testing these concepts has been possible through correlative studies and using paleontological records. Birds and primates are particularly interested in this context because they independently evolved relatively larger brains than other vertebrates (Jerison 1973). For example, in birds foraging innovations have been reported in over 800 species and include a wide range of behaviors, from eating a novel food to using tools. Overington et al.' results suggest that the cognitive behavior required to

perform a wide variety of novel foraging techniques supports the positive relationship between innovativeness and brain size in birds (Overington et al. 2009). Deaner and colleagues found that absolute brain size is a good predictor of a global cognition index extracted from meta-analyses in primates. Moreover, there was no indication that neocortex-based measures were superior to measures based on the whole brain (Deaner et al. 2007). Using the fossil and extant dataset, researchers (Smaers et al. 2021) found that shifts in allometric slope underpin major transitions in mammalian evolution and are often primarily characterized by marked changes in body size. These results reveal that the largest-brained mammals achieved sizeable relative brain sizes by highly divergent paths. In addition, Sol et al. (Sol et al. 2022) estimated neuron numbers in 111 bird species. They showed that the number of neurons in the pallial telencephalon is positively associated with intelligence.

However, there are some essential problems with comparative data; cognitive performance is challenging to measure on a broad scale, identified correlations do not necessarily imply causative relationships, and, finally, they do not explain why the brain should be bigger to perform a particular behavior. There is also ample literature demonstrating that brain size alone is a poor proxy for CA (Healy and Rowe 2007; 2013). Comparative studies are weakened by inconsistencies in the methodology of behavioral assays and data collection. For example, Hooper et al. (Hooper, Brett, and Thornton 2022) investigated variation in brain and body size measurements across >1000 bird species. They found an extensive variation in brain and body size across source datasets, resulting in inconsistent conclusions drawn from comparative brain size models based on the particular data source. In particular, using a subset of data on Corvids, the authors showed that depending on the data set used in the analysis, multiple contradictory conclusions can be drawn about the principal factors affecting brain size evolution.

The evolution of brain size and CA were further investigated by Chambers et al. (Chambers, Heldstab, and O'Hara 2021), who tested the effect of social, ecological, and life-history traits on variation in encephalization and specific brain regions amongst extant primates and carnivores. In primates, researchers identified significant associations between brain size, diet, and sociability. Conversely, the association between brain size and sociability was insignificant in carnivores. In both taxa, however, the study revealed complex associations between metabolic costs of maintenance of brain tissue and extended developmental periods, reduced fertility, and maximum lifespan (Chambers, Heldstab, and O'Hara 2021). Together, these studies suggest that differences in methods of brain measurement, particularly differences in

measurements of brain regions, should be considered as potential confounding factors affecting broad-scale comparative studies on brain size and CA.

Apart from above discussed methodological difficulties inherent to comparative studies on brain size vs. CA, they are all hampered by the lack of a robust association between the two, as exemplified by a simple comparison of brain size and CA of an elephant and a man, the latter having three times smaller brain, but apparently superior cognitive abilities (Herculano-Houzel et al. 2014). The weakness of a principle, 'The bigger brain is smarter' is further exposed by comparing the social skills of ravens and apes (Pika et al. 2020). The comparison indicates that the lack of a specific cortical architecture (absent in ravens but present in apes) does not hamper the elaborate cognitive skills of this bird species.

4.5 Testing the links between energetics and the evolution of brain size and function

As discussed above, the evolution of the energetics and brain size and function, and their interrelation, are difficult to test because the underlying physiological mechanisms are not preserved in the paleontological record. The evolutionary plausibility of the hypotheses discussed in Section 4.2 can thus only be tested if they reflect more general evolutionary principles applicable to extant animals characterized by a positive association between enlarged brains and enhanced cognitive abilities. As discussed in Section 4.3, most studies on the associations between brain size and energetics, internal organs, and CA were correlative and confounded by potential methodological weaknesses inherent to comparative analyses carried out on the between-species level (Van Schaik et al. 2021). Most importantly, natural selection acts within-not between- species. Thus, although comparative studies are informative for broad evolutionary patterns, they do not allow for the demonstration of causative links (Harvey and Pagel 1992), which is essential for solid tests of hypotheses such as the ET and EB.

Even though a direct re-creation of physiological traits of extinct species is impossible, the evolutionary processes leading to encephalization can be at least partly emulated through artificial selection experiments carried out on extant species (Garland 2003). Such experiments offer a powerful tool for investigating the linkages between anatomical and physiological traits that influence physiological performance and testing how performance capacities may constrain or facilitate behavioral evolution (Swallow et al. 2009). The artificial selection experiment, purposely developed to test the ET hypothesis has been carried out by Kotrschal et al. (Kotrschal et al. 2013). They selected lines of guppies (*Poecilia reticulata*) for relative brain

size and analyzed the between-line variation of the gut size and CA using a 'Numerical Learning' test. Kotrschal et al. (2013) demonstrated that large-brained females performed better in a numerical learning assay than small-brained females. Moreover, in a direct support of the ET hypothesis, large-brained lines had smaller guts at the expense of producing fewer offspring. Although Kotrschal et al.'s experiment is insightful, it is also hampered with two shortcomings. First, their demonstration of the cognitive advantage of having a larger brain has turned out to be questionable (Healy and Rowe 2013). More importantly, however, life history and physiology of fish is so far removed from that of homeotherms, that it bears unclear relationship with possible selection for encephalization in mammals, particularly in primates.

Here, I tested hypotheses outlined in Section 4.2 in a mammalian model of experimental evolution for the first time. I used line types of laboratory mice subjected to artificial selection on high (H-) or low (L-) basal (BMR), or high maximum (VO_{2max}) metabolic rates (PMR) (Książek, Konarzewski, and Łapo 2004; Gębczyński and Konarzewski 2009; Joanna Sadowska, Gębczyński, and Konarzewski 2017), the traits widely accepted as pre-requisites for the evolution of homeothermy and large brain size (A. F. Bennett and Ruben 1979; Lovegrove 2017). Line types of mice divergently selected for BMR are characterized by a conspicuous 40% difference between low (L-BMR) and high (H-BMR) line types, derived from Swiss Webster outbred strain of laboratory house mice (Swallow et al. 2009). Selection for high maximum (VO_{2max}) is carried out on four replicated lines and four control (random bred) ones. Importantly, in the course of selection, both BMR and VO_{2max} are corrected for body mass effect.

Earlier studies showed that selection for VO_{2max} did not affect BMR, while selection for BMR resulted in a reduction of PMR (Gębczyński and Konarzewski 2009). Therefore, the analyses of the between-line type differentiation of the energetic costs of maintenance of the brain, internal organ masses, and the size vs. variation in BMR and PMR will allow for the cross-tests of the predictions of the ET and EB hypothesis. My experimental model allowed me to analyze not only trade-offs between anatomic and physiological traits, but also directly test the directionality of associations between energy expenditures and the rate of learning through a battery of carefully controlled behavioral tests. Finally, the analysis of the between-line variation in neurophysiology and cell architecture of the hippocampus, the brain structure underlying learning in the performed behavioral tests, allowed me to identify the underpinnings of the differences in CA related to the hypotheses in question.

5 Aims and objectives

In my thesis, I posed a general question: How were the energetic costs of enlarged mammalian brains overcome? To answer this question, I tested predictions of two major hypotheses: the ET and EB, presented in Section 4.2 of the Introduction. To do so, I used line types of laboratory mice subjected to artificial selection on high (H-) or low (L-) basal (BMR), or high maximum (VO_{2max}) metabolic rates (PMR), the traits widely accepted as pre-requisites for the evolution of large brain size. This experimental model allowed me to analyze (1) trade-offs between brain size, gut size, and BMR or PMR, (2) test the directionality of associations between BMR or PMR and the rate of learning (a proxy of cognitive abilities, CA), (3) analyze the size and cell architecture of the hippocampus, a brain structure critically involved in the formation, organization, and retrieval of new memories, that most likely underlie the differences in CA. Then I compared the results of my studies with the predictions of the ET and EB hypotheses.

More specifically, my thesis aimed at:

(1) testing the existence of association between BMR or PMR and cognitive abilities (CA) and brain size. Such associations are fundamental for all hypotheses relating the evolution of brain size to energy expenditures.

(2) testing the existence of the brain-gut trade-off. This trade-off constitutes the central prediction of the ET hypothesis. Conversely, the EB hypothesis does not predict the brain-gut trade-off and points to the positive association between metabolic rates and brain size.

(3) identifying cellular underpinnings of the between line type differences in CA, concentrating on the hippocampal cell size and function. In particular, I tested the among-line type differences in:

a) synaptic plasticity, which gave an insight into neuronal mechanisms underlying the observed variation in learning abilities;
b) the size and density of hippocampal neurons and astrocytes, to test whether higher CA are associated with more neurons;
c) the number of dendritic spines in key areas of the hippocampus, as a proxy of neuronal connectivity;
d) the proportion of astrocytes feeding hippocampal neurons since the number and size of neurons is not the only determinant of brain information-processing capacity;

e) the differences in the density of cytochrome C oxidase (CCO) in the hippocampus to estimate the metabolic costs of its activity.

The results are presented below as two consecutive experiments.

The results descried in the first part of the thesis were published in Proceedings of the Royal Society B:

Goncerzewicz, A., Górkiewicz, T., Dzik, J. M., Jędrzejewska-Szmek, J., Knapska, E., & Konarzewski, M. (2022). Brain size, gut size and cognitive abilities: the energy trade-offs tested in artificial selection experiment., *Proceedings of the Royal Society B* 289(1972). https://doi.org/10.1098/RSPB.2021.2747

6 Brain size, gut size and cognitive abilities: the energy trade-offs tested in an artificial selection experiment

6.1 Materials & Methods

6.1.1 Animals

I used 3-4 month-old female mice from two concurrent selection experiments carried out at the Faculty of Biology, University of Bialystok. The first experiment consisted of two line types of mice divergently selected for high (H-BMR) or low (L-BMR) body-mass-corrected Basal Metabolic Rate (BMR), quantified according to the measurement procedure outlined below. The resulting divergence between those two non-replicated line types is sufficiently large to be confidently attributed to the applied selection rather than to genetic drift (see section 6.1.5.3). I also used female mice from the second selection experiment, in which eight genetically isolated Swiss-Webster laboratory mouse lines were established. In four of the lines, forming the Peak Metabolic Rate (PMR) line type, mice were selected for VO_{2max} quantified as the highest body-mass-corrected oxygen consumption averaged over 2 min of a 5 min swim in a 25 °C water (Gębczyński and Konarzewski 2009; Joanna Sadowska, Gębczyński, and Konarzewski 2017). The other four lines form the random bred (RB), control line type (Gębczyński & Konarzewski, 2009; Sadowska et al., 2017).

In selection experiments animals were housed in same-sex and same-family groups of four to five per cage (365 x 207 x 140 mm) at 23°C and 12:12 light schedule. Throughout the experiment animals were fed a standard diet (12.8 kJ g−1 metabolizable energy, 17.0 kJ g−1 caloric value manufactured by Labofeed, Kcynia, Poland.

I treated animals following the ethical standards of the European Union and Polish regulations. All procedures were approved by the Local Ethical Committee on Testing Animals.

Measurements of BMR, organ masses, and behavioral tests were carried out on separate groups of animals randomly drawn from a stock of animals not qualified for further selection. Each mouse was randomly selected from a separate family belonging to a particular line. The numbers of animals used in particular trials are given in Table 1.

Table 1. The number of animals included in the statistical analyses (n) and generation of origin (F). N.M.- not measured.

Trait measured	Selection direction							
	RB		H-BMR		L-BMR		PMR	
	n	F	n	F	n	F	n	F
BMR	40	37	76	51	74	51	35	37
Body Mass	40	37	35	51	34	51	39	37
Liver Mass	39	37	35	51	34	51	39	37
Heart Mass	40	37	35	51	34	51	39	37
Kidneys Mass	40	37	35	51	34	51	39	37
Brain Mass	40	37	35	51	34	51	39	37
Correct Nosepokes Sucrose	89	38, 42, 43	19	52, 53	21	52, 53	71	38, 42, 43
Incorrect Nosepokes Quinine	27	46, 47	13	68, 69	10	68, 69	33	46, 47
Locomotor Activity	95	38, 42, 43	15	52, 53	21	52, 53	78	38, 42, 43
Number of Licks to the Bottle with Sucrose	90	38, 42, 43	18	52, 53	20	52, 53	69	38, 42, 43
Fear Extinction	36	42, 43	20	53	14	53	37	42, 43

6.1.2 Measurements of Basal Metabolic Rate

I used two separate positive pressure open-circuit respirometry systems fitted to two Sable Systems FC-1B oxygen analyzers. In each system, the outside atmospheric air was dried and then forced through a copper coil submerged along with metabolic chambers (each 350 cc in volume) in a water bath stabilized at 32°C (a temperature that is within our animals' thermoneutral zone) to equalize and control the temperature. The air stream was then divided

to four independent streams (including the reference stream); each fed at 400 mL min^{-1} to a separate mass flow controller (Sierra Instruments, Monterey, CA or ERG1000, Warsaw, Poland), individually calibrated with a bubble flow meter (Optiflow 420, Supelco, Bellefonte, PA). The air was then forced through individual metabolic chambers, and further through a computer-controlled channel multiplexer (Sable Systems, Las Vegas, NV). Air was after that scrubbed of CO_2, dried one more time, subsampled at the rate of 75 mL min^{-1}, and fed to an oxygen analyzer. BMR (STPD) was calculated with Wither's equation (Withers 1977) and defined as the lowest stable reading that did not vary by more than 0.01% oxygen concentration for at least 4 min.

All metabolic trials were carried out between 0800 and 2000 hours. Before measurements, mice fasted for 6 h, then weighed to the nearest 0.1 g. During the two h measurement period, each metabolic chamber was sequentially monitored for about 20 min, followed by a 2-3 min break for zeroing.

6.1.3 Behavioral tests

6.1.3.1 Measurements of cognitive abilities in the IntelliCage System

The IC system is a fully automated, computer-controlled system for the behavioral assessment of mice that live in social groups (Lipp et al., 2005; Lipp, 2005). This system is suitable for long-term monitoring cognitive functioning in group-housed mice (Figure 1c).

The IC consists of a large standard rat-size cage 20.5 cm high, 40 cm × 58 cm at the top, and 55 cm × 37.5 cm at the base. The cage is equipped with four operant learning chambers fitted into the corners of the housing cage. Access into the chamber is only possible through a tube with a built-in transponder codes reader (antenna) that restricts access to the learning chamber to only a single mouse at a time. Each corner, equipped with a proximity sensor, contains two openings permitting access to drinking bottles. An automatically operated door controls access to the liquid. Poking a nose into the openings (nose poke response) activates an infra-red beam-break response detector. Each visit to the operant chamber, each nose poke and the amount of water consumed (number and duration of licks) is recorded for each animal. The cage control unit permits access to particular bottles according to schedules individually pre-programmed for each mouse. The cage is equipped with a sleeping shelter in the center, with a feeder on its top providing food ad libitum.

A week before the experiment, the mice were sedated with isoflurane and injected with a glass-covered microtransponder (11.5mm length, 2.2mm diameter; DataMars) with a unique code recognized by sensors installed in the learning chambers (Masuda, Kobayashi, and Itohara

2018). After the transponder implantation procedure, subjects were moved from the housing facilities to the experimental rooms. The animals were then transferred to the IC, each housing 10-12 mice randomly drawn from the stocks of their parental lines. Individuals from the same line type were housed together to minimize the possible effect of social context (Kiryk et al. 2011). The number of mice living in the cage was adjusted to exclude competition for access to the bottles (Knapska et al. 2013). Mice housed in each IC were maintained in a 12:12 light schedule (same as the maintenance conditions in their home animal facility) and subjected to learning tasks. The mice were not disturbed except for the technical breaks and cage exchange (once a week). Apart from this, the subjects did not have contact with an experimenter. The behavioral activity of mice was recorded 24h a day throughout the adaptation period and testing procedures and was monitored via remote desktop software (TeamViewer software). Before experiments were carried out in the IC system, mice were not water- and food- deprived.

6.1.3.2 Appetitive discrimination protocol

Cohorts of mice were subjected to the 15-day IC protocol (Figure 1a), consisting of four phases: simple adaptation, nose poke adaptation, place preference learning, and reward-seeking discrimination learning with each of 8 bottles containing tap water and the assessment of reward discrimination learning (5 days) with highly motivating reward (10% sucrose solution) accessible in one of the bottles. During the simple adaptation phase lasting 4 days, doors in all conditioning units (corners) were open and access to water was unrestricted. During the following three days of nose poke adaptation all doors were closed and opened only when an animal put its snout (nose poke response) into one of the two openings. When an animal removed its nose from the opening, the door closed automatically. During the place preference learning phase (days 8-10) access to the drinking bottles was restricted to only one of the IC learning chambers for no more than 3 mice drinking from the same conditioning unit. During discrimination learning (days 11-15) mice had a choice between nose poking (operant response) to the bottle containing tap water or to the bottle containing a reward (sweetened water) placed in the same conditioning chamber. They had to remember the location of the reward to perform the correct response. A total number of nose pokes, a number of visits (activity) and liquids consumption were recorded.

6.1.3.3 Aversive discrimination protocol

In an aversive learning task, a group of naive mice was subjected to the above-described reward-seeking discrimination learning procedure extended by three additional phases (Figure 1b). In the first one mice had access to water in all four corners for 2 days (days 16-17). During the next phase, lasting 2 days, mice had access to bottles only in one of the corners, which was different from the corner with the reward in the reward-seeking discrimination training. In the third phase, lasting 5 days, the bottle preferred during the previous two days was replaced with a bottle containing 0.005 M quinine solution evoking aversive, bitter taste perception in mice of all studied line types. Changes in the number of nose pokes to the bottles containing quinine (i.e., incorrect responses) recorded during the first critical 24 h were considered a response.

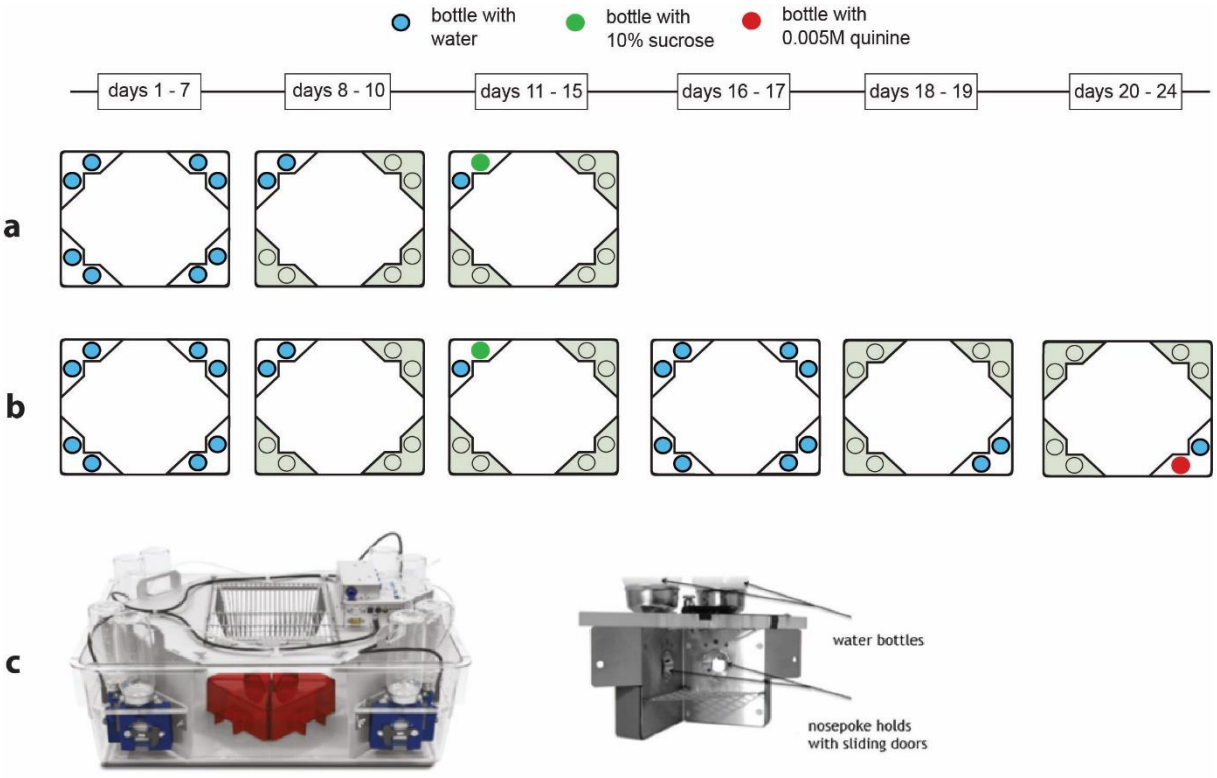

Figure 1 Scheme of the IC experiments testing cognitive functioning.

In the reward-seeking discrimination learning task (a) naïve mice were subjected to experimental procedures that consisted of simple adaptation phase (days 1-4, not shown), nosepoke adaptation phase (days 5-7), place preference learning (days 8-10) and reward-seeking discrimination learning (reward: 10% sucrose solution, days 11-15). These phases were repeated in the aversive discrimination task (b), in which another group of naïve mice was also subjected to additional phases: the nosepoke adaptation (days 16-17) and place preference learning to a different corner (days 18-19). Next, on days 20-24, mice were exposed to an aversive discrimination learning procedure with a 0.005 M quinine solution; (c) IC, and automated corner arrangement.

The IC system recorded the number of visits, nose pokes, and tube licks in 12 h time intervals. PyMICE then assembled all raw data (Python library for mice behavioral data analysis, (Kowalski et al., 2015). A critical part of this data set consisted of the last 24 h of the place preference learning phase and the next 24 h of the reward-seeking or aversive discrimination learning phases and was used in the further analysis (Knapska et al. 2013). Data were excluded from the analysis when an animal did not learn to obtain water during the preparatory place learning phase and thus was unable to correctly perform subsequent reward discrimination task, or, in rare cases, when the animal's locomotor activity was extremely low (as reported by the IC software).

6.1.3.4 Fear conditioning

The mice were subjected to Pavlovian contextual fear conditioning in a fear conditioning chamber (MED ASSOCIATES). The training was carried out according to a classic paradigm (Curz, Rustay, and Browman 2009). It consisted of 3 min adaptation period and 5 footshocks lasting 1 s and having 0.6 mA intensity, which were applied with interstimulus intervals of 2 min. The animals were removed from the experimental cage to their home cages 2 min after the last footshock was applied. The next day the animals were tested in the same cage for 3 min. Fear of the context was assessed by measuring freezing behavior. To avoid counting brief inactivity as freezing, freezing was scored only if the mouse was immobile for at least 1 s. The freezing observations were transformed to a percentage of total observations in 30 s intervals. The experiment was performed in the context that included: transport cages carried by an experimenter, bedding in the transport cage, white light on and air-conditioning on in the experimental room, conditioning chamber with a metal grid floor, fan, and white light on.

6.1.4 Morphometrics

Animals subjected to the reward-seeking discrimination learning trial were killed by cervical dislocation and dissected. Brain, heart, liver, and kidneys were excised, blotted from excess fluids, and weighed to an accuracy of 0.001 g.

6.1.5 Statistical analysis

6.1.5.1 Data handling

The NORMAL, ROBUSTSCALE and PLOT options in the PROC UNIVARIATE of SAS were used to check the assumption of normality and homoscedasticity of the data and their residuals. I log10 transformed BMR measurements, because of a strong right-skewness of the distribution. Inspection of histograms of residuals and quantile plots of masses of organs revealed 6 outlying values, which were excluded from analyses of the respective organs, but were retained in analyses of other organs, in which their residuals did not stand out. Offending values are marked in red in a data set deposited in DRYAD (https://doi.org/10.5061/dryad.bk3j9kd78). Furthermore, 5 individual organ masses were declared as missing, because of mistakes in data recording. In behavioral trials some individual observations are also declared as missing due to uncertainty of those observations related to temporary malfunction of the sensors (resulting in, e.g., negative counts of activity of individual animals). Descriptive statistics of anatomical and physiological traits are presented in Table 2.

Table 2. Descriptive statistics of anatomical and physiological traits: arithmetic means and standard deviations.

Trait measured	**Selection direction**							
	RB		**H-BMR**		**L-BMR**		**PMR**	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
BMR (ml O_2/h)	52.2	7.0	64.4	7.8	40.1	5.2	53.4	7.5
Body Mass (g)	29.1	1.8	31.1	2.5	29.8	2.9	28.2	2.3
Liver Mass (g)	1.50	0.18	1.84	0.25	1.14	0.16	1.46	0.20
Heart Mass (g)	0.15	0.01	0.20	0.05	0.14	0.02	0.17	0.02
Kidneys Mass (g)	0.36	0.04	0.45	0.04	0.32	0.04	0.39	0.05
Brain Mass (g)	0.46	0.02	0.49	0.04	0.46	0.04	0.45	0.03

6.1.5.2 Mixed ANCOVA Models

All statistical analyses were carried out by a mixed model extension of a general linear model (Mixed procedure of SAS/STAT® 14.1 User's Guide). REML method of estimation and variance components constrained to non-negative values were used. Repeated measures mixed ANCOVA models were fitted with either equal or unequal residual covariance (with the TYPE option of REPEATED statement of procedure Mixed set to CS (compound symmetry) or Toep

(Toeplitz structure). For the final analysis, I chose the covariance structure yielding a lower Akaike information criterion (AIC).

Data on BMR and masses of internal organs were analyzed with ANCOVA with line type affiliation as a fixed factor, body mass as a covariate, and the Line type × Body mass interaction. Initial BMR analyses also included the respirometric system and metabolic chamber coded as fixed factors. Their effects (as well as Line type × Body mass interaction) were never significant ($p > 0.05$) and therefore were dropped from final analyses.

Repeated measures analysis of covariance (ANCOVA) was used to analyze the among-line type differences in total numbers of visits to all four corners summed over four continuous 12 h (dark followed by light) periods of observation (thereafter Period), covering the last 24 h (thereafter Day) of place preference learning, and first 24 h of reward-seeking discrimination learning.

In the reward-seeking, and aversive discrimination learning tasks, I analyzed the number of nosepokes to the bottles in a corner assigned to a given animal during a critical 48 h of trials. During the first two 12 h Periods (located left of the vertical dashed line denoting the timeline of the experiment, Figure 2 A-C), both bottles in the corner contained water. Subsequently, at the onset of the following two 12 h Periods one bottle was filled with 10% sucrose (reward-seeking discrimination learning) or 0.005M quinine solution (aversively motivated discrimination learning). In the ANCOVA model, the numbers of correct responses (i.e., nose pokes to the bottle with sucrose) or incorrect responses (in the case of quinine) were corrected for (1) the dark and the light experimental Periods and the respective Day (effects of both Period and Day coded as a fixed factor of a factorial design) (2) number of nose pokes to the bottle with tap water located in the same corner coded as a covariate. I used an analogously structured model to analyze the number of licks of the bottles containing tap or sweetened water.

To analyze the rate of changes in freezing response, I used ANCOVA with the line type as the main factor, individual identification of animals, and line (nested within line type) as random factors with time (subsequent 30 s intervals) as a covariate.

In all analyses, replicated lines were nested within line types as the random factor of the model (4 replications in the RB and PMR line types, respectively, but 1 line for H-BMR and L-BMR line types, respectively, as they were not replicated; 10 lines in total). The respective error mean square for 10 lines was used as the denominator of the F statistics testing the effect of line type affiliation. Hence, the df for the among line type comparisons was 3 (for the F numerator) and 6 (for the denominator).

The significance of the random effect of lines nested within line types was tested with the likelihood ratio (LR) test, which employs $\chi 2$ distribution of -2 times the log likelihood of fit of

the model of a given dependent variable with only the fixed effects minus -2 times the log likelihood from the full model (with random factor). The $\chi 2$ values with their respective p values are presented in Table 3.

Table 3. Likelihood ratio test (LR) statistics for the significance of the random effect of replicate lines (nested within line type) in ANCOVA mixed models of analyzed traits.

	Replicate line effect			
Dependent variable	**s^2**	**SE**	**χ^2**	**p**
BMR (ml O_2/h)	0.000926	0.001239	1.1	>0.1
Body Mass (g)	1.27	1.04	6.4	<0.02
Liver Mass (g)	0.0062	0.0059	3.3	>0.05
Heart Mass (g)	N.E.			
Kidneys Mass (g)	0.00050	0.00037	11.1	<0.001
Brain Mass (g)	0.00030	0.00021	14.9	<0.001
Correct Nose pokes sucrose	6.73	5.84	5.3	<0.025
Activity	7.96	10.73	1.2	>0.2
Licks	113856	82518	12.8	<0.001
Incorrect Nose pokes quinine	19.2711	19.5530	2.9	>0.05

Statistical models initially included all respective interactions. The models were then step-wise reduced by removing non-significant interactions ($p > 0.05$) except those that were considered integral components of the tested hypotheses (e.g., Line type × Period interaction in statistical models of behavioral tests). Initial models of behavioral tests also included the effects of generation and batch of animals (a group of animals simultaneously subjected to the behavioral test, coded as fixed factors). The generation effect was never significant ($p > 0.05$), and therefore was dropped from the final analyses. However, the batch effect was significant in several analyses, as indicated in Table 3.

As stated in Section 4.5 of the Introduction, I tested for '(1) the existence of the brain-gut trade-off and (2) positive associations between BMR or PMR and CA and brain size'. I started with

testing (1) and identified the H-BMR and L-BMR line types with the most divergent BMRs and masses of internal organs relevant to testing the gut-brain trade-off. Therefore, in the behavioral part of our study, I concentrated on comparing the H-BMR and L-BMR line types with the remaining line types, by means of the custom made a priori contrasts of the H-BMR or the L-BMR line type vs. the rest. Those planned custom hypotheses were tested for the differences in slopes of the changes in the number of nose pokes as depicted in Figure 5. They were carried out with the Contrast statements of the procedure Mixed of SAS (for details, see SAS script provided in Supplementary Material).

To illustrate the use of the contrast custom hypotheses, let me consider the Line type × Time interaction of the extinction of freezing response to a perceived threat illustrated in Figure 6. I first tested the heterogeneity of slopes of changes in freezing response of the RB, H-BMR, and PMR line types. The resulting test statistic indicated that the slopes did not differ ($F_{2,530}$ = 1.51; p = 0.22). In the second custom test, I asked whether the slope of the L-BMR line type differed from the remaining line types. This difference was statistically significant at p = 0.005 with $F_{2,530}$ = 7.79. Thus, taken together, both custom tests suggested that the statistical significance of the Line type × Time interaction in the ANCOVA analysis of the extinction of freezing response can be attributed to the difference between the slope of the rate of this extinction in the L-BMR mice vs. the slopes of other line types.

6.1.5.3 Evaluation of the effect of genetic drift in divergent selection for BMR

Because of time and resource limitations, divergent selection for BMR is not replicated. This may confound the interpretation of my findings, because one cannot rule out the possibility that the between-line type differences between the L-BMR and H-BMR mice were due to genetic drift rather than direct effects of the applied artificial selection (Henderson 1997; Konarzewski 2005). To evaluate the potential effect of genetic drift, the L-BMR vs. H-BMR differences in internal organ masses were analyzed according to the guidelines recommended by Henderson (Henderson 1997) and Konarzewski et al. (Konarzewski 2005). Briefly, first the within-line standard deviations (SDs) of analyzed traits were calculated. Since individual mice used in the measurements came from different families, those SDs can be interpreted as phenotypic SDs, being a square root of the product of narrow sense heritability of a given trait and its genetic SD (Henderson 1997). Then, thus obtained SDs were used to calculate standard deviation *SDx* weighted across both selected line types were used:

$$SD_x \cong \sqrt{\frac{SD^2_{H-BMR}(n_{H-BMR}-1)+SD^2_{L-BMR}(n_{L-BMR}-1)}{n_{H-BMR}+n_{L-BMR}-2}}$$

where $SD_{H\text{-}BMR}$ and SD_{L_BMR} are phenotypic standard deviations of the studied trait *x*, for the H-BMR and L-BMR line types, respectively, and $n_{H\text{-}BMR}$ and $n_{lH\text{-}BMR}$ are the numbers of families in the respective line types (Konarzewski 2005; Henderson 1997).

Next, following Henderson (Henderson 1997), the magnitude of separation of the line types for the studied trait - were expressed as the multiples of the *SDx*:

$$d_x = \frac{\bar{x}_{H-BMR} - \bar{x}_{L-BMR}}{SD_x},$$

where $\bar{x}_{H-BMR}$ and $\bar{x}_{L-BMR}$ are within line type means of the studied trait.

Finally, I compared thus calculated d_x with the confidence limits of the magnitude of the between line type separation expected under genetic drift, calculated according to equation 16 from Henderson (Henderson 1997):

$$\sigma_{(dx)} \cong 2\sqrt{(h^2{}_x F + 1/n)},$$

where h^2 is the narrow sense heritability of a given trait, *F* is the averaged coefficient of inbreeding of the selected line type (Hill and Mackay 2004), and *n* is the number of families subjected to selection in both line types.

Values of d_x falling below the upper boundary of the 95% confidence interval (thereafter d_{drift}) should be ascribed to the effect of genetic drift and sampling error alone, acting in the absence of genetic correlation between the primary selected trait (BMR) and other analyzed traits. Conversely, values of d_x exceeding that of d_{drift} indicate genetic correlations (Henderson 1997). Depending on the actual demand, between 40 and 80 families within each selected line type were maintained. For calculations, I therefore assumed a conservative number of families equal to $n = 40 + 40$ for both line types. The coefficient of inbreeding (*F*) for generations F51, F52, and F53 used in my study were 0.29, 0.30, and 0.31, respectively. I used $n = 80$, and averaged value $F = 0.30$ to calculate d_{drift} for different values of h^2, and compared them graphically (Figure 4) with the d_x computed for organ masses.

6.2 Results

6.2.1 Morphometrics

6.2.1.1 Mixed ANCOVA analyses

Mean body mass did not differ among mice from the four line types (Table 4). H-BMR mice were characterized by conspicuously higher body mass-corrected BMR then mice of all other line types (Table 4, Figure 2A and Figure 3A). Their metabolically expensive internal organs (liver, heart and kidneys) were also larger than in mice of other line types (Figure 2B, C, D and Figure 3B, C, D). Yet, their brains were not significantly larger, particularly when compared with their direct counterparts- the mice from the L-BMR line type (Figure 2E and Figure 3E).

Table 4. ANCOVA results for BMR and organ masses.

	Line type	**Body Mass**
Body Mass	$F_{3,6} = 1.69$ $p = 0.27$	-
BMR	$F_{3,6} = 38.05$ $p < 0.001$	$F_{1,212} = 82.19$ $p < 0.001$
Brain	$F_{3,6} = 0.70$ $p = 0.58$	$F_{1,132} = 4.09$ $p = 0.04$
Liver	$F_{3,6} = 10.18$ $p = 0.009$	$F_{1,132} = 28.56$ $p < 0.001$
Heart	$F_{3,6} = 29.93$ $p < 0.001$	$F_{1,136} = 13.91$ $p < 0.001$
Kidneys	$F_{3,6} = 6.34$ $p = 0.03$	$F_{1,135} = 48.85$ $p < 0.001$

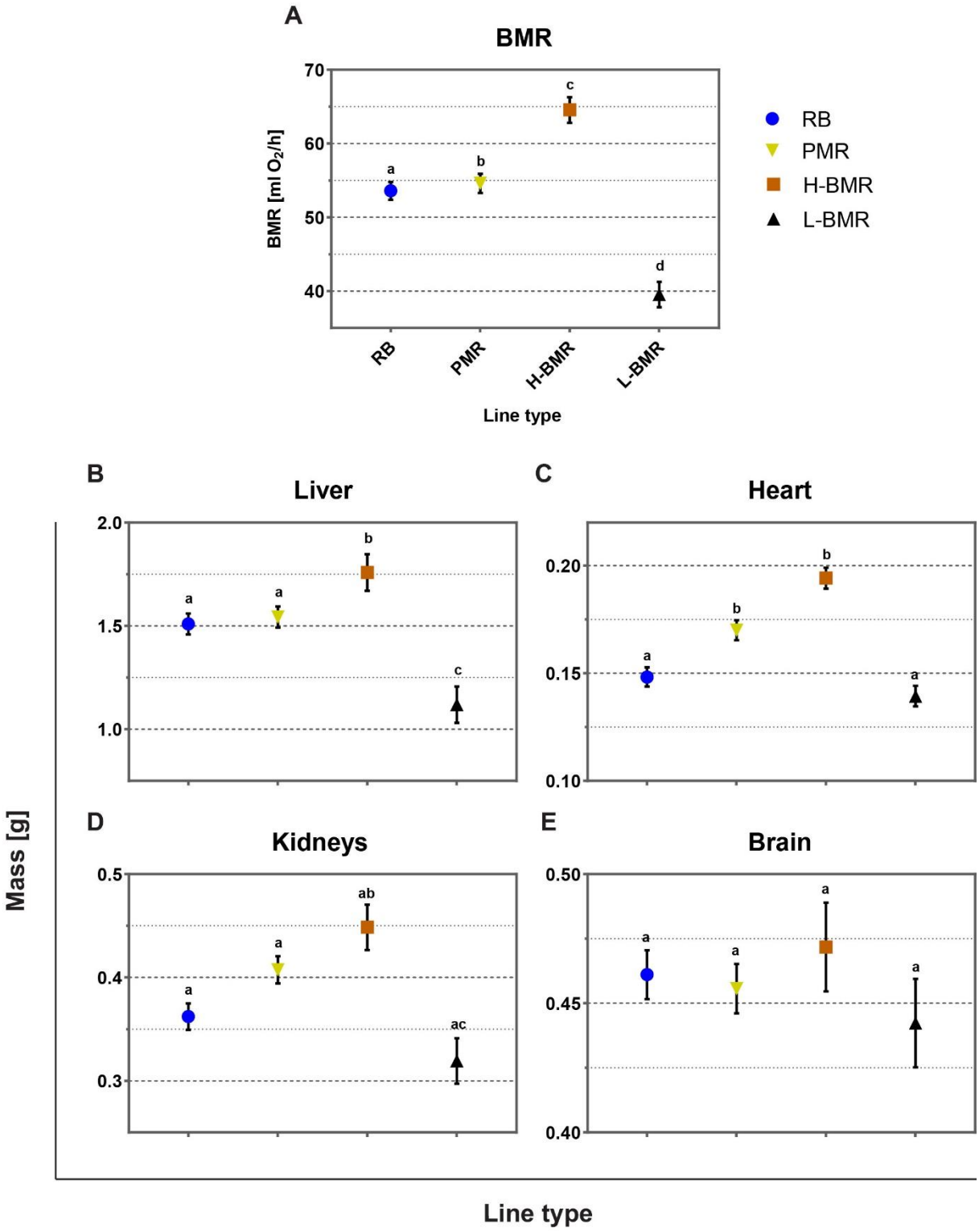

Figure 2. (A) Basal Metabolic Rate (BMR) in line types of mice selected for high (H-BMR) or low (L-BMR) Basal Metabolic Rate (BMR), Peak Metabolic Rate (PMR aka VO_{2max}). RB-random-bred line type; (B-D) masses of internal organs, and (E) brain mass. Values are body mass adjusted means with standard errors calculated from ANCOVA. Figure bars labeled with different letters differ significantly from each other at p = 0.05

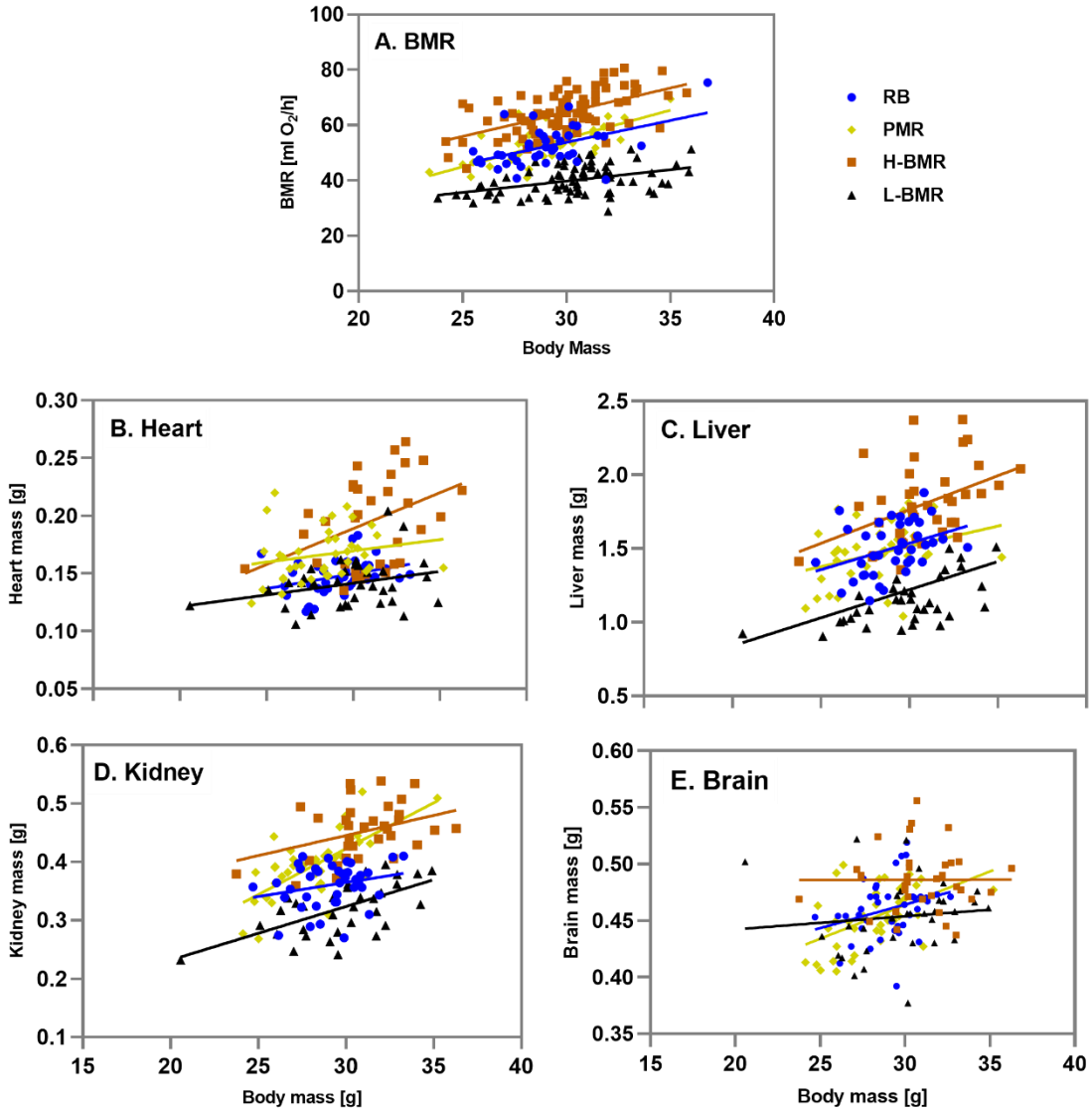

Figure 3. (A) Individual data for Basal Metabolic Rate (BMR), masses of internal organs (B-D), and brain mass (E) plotted against body mass in mice of studied line types. The slopes of the regression lines did not differ significantly between the line types at p = 0.05.

6.2.1.2 Evaluation of the effect of genetic drift on organ masses

Inspection of Figure 4 reveals that the d_xs of internal organ masses (heart, liver, and kidneys) are above the value of d_{drift} in the whole range of h^2. Conversely, the *dx* of the brain mass falls well below d_{drift} for h^2 = 0.64 - published estimates of h^2 of brain mass in laboratory mice (Atchley et al. 1984). Thus, a weak separation of the H-BMR and L-BMR line types for brain mass does not allow for its firm recognition as a correlated response to selection on BMR, rather than the effect of genetic drift (Atchley et al. 1984). In contrast, the analysis suggests that the between line type differences in internal organ masses (heart, liver, and kidneys) are genuine effects of the applied selection regimen and indicate their positive genetic correlations with BMR.

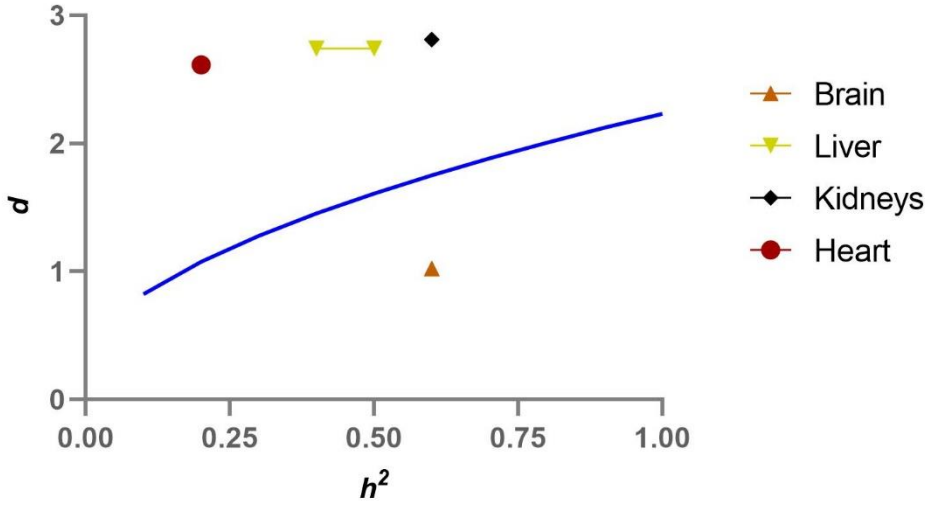

Figure 4. The magnitude of separation of the H-BMR and L-BMR line types for the studied traits (d) along with the upper boundary of the 95% confidence interval (d_{drift} , blue line) plotted against the narrow sense heritability (h^2). The d values for organ masses were located according to published values of h^2 (heart- [8]; liver-[8, 9]; kidneys-[10]; brain mass- [7]). Unlike the d values for other traits, the d for the brain mass falls below d_{drift} and, therefore, should be ascribed to the effect of genetic drift.

Thus, both a direct comparison of all line types presented in section 6.1.5.3 and an evaluation of H-BMR vs. L-BMR difference for the effect of genetic drift indicated that the brain-gut trade-off as well a positive genetic correlation between BMR and brain size were not observed.

6.2.2 Behavioral tests

To compare learning abilities, I trained mice in the IC automated system that allows for individual activity assessment and learning of group-housed mice (Knapska et al. 2013). In an initial acclimatization period, mice were able to access water in any of the four corners of the IC. Each corner had two separate bottles with tap water that the mouse could choose between.

During the place preference learning, water access for each mouse was restricted to one of the four corners. Next, in the reward-seeking discrimination learning, one of the bottles was filled with a reward- 10% sucrose solution and the learning progress was scored as the number of nose pokes that opened access to the bottle with reward (correct responses).

In comparison to the last 24 h of the previous training phase of the training, during the next 24 h, all mice increased the number of nose pokes to the bottle that now contained the reward. However, H-BMR mice accessed the reward more often than their L-BMR, PMR and random-bred counterparts (Table 5, Figure 5A). Most importantly, a highly significant Line-type × Time interaction indicated that the H-BMR mice learned the rewarded response faster than the other animals ($F_{1,761} = 15.1$; $p < 0.001$ for the planned comparison of the slope of change of the number of nose pokes in the H-BMR mice vs. other line types).

To test whether the improved learning could be attributed to changes in thirst or taste discrimination, the number of licks from the bottles that contained sucrose solution was analyzed. I did not observe any differences among the line types in the amount of sweetened water consumed (Table 5). Further, because differences in general activity could potentially influence the obtained results, I compared the numbers of visits to all corners during the reward-seeking discrimination learning and the adaptation phase. The rate of visiting corners did not differ among the line types (Table 5), excluding the possibility that changes in general activity could explain the differences in learning.

To exclude the possibility that the superior learning response of the H-BMR mice was solely limited to the reward-seeking context or higher motivation to perform a nosepoke response, I subjected another group of naive mice to a separate behavioral test in the IC system as described above, but with the reward-seeking discrimination learning followed by an aversive cue discrimination task. In this additional task mice learned to avoid an aversive cue from a water solution of 0.005 M quinine placed in one of the IC corners. In the reward-seeking part of the trial, the H-BMR mice again accessed the reward more often than the mice of other line types (Line-type × Day interaction, $F_{3,306} = 7.13$; $p < 0.001$). In contrast, when sucrose solution was replaced with quinine, the number of nose pokes to the bottles now containing its solution decreased in the H-BMR mice and remained unchanged in other line types (Figure 5B). This resulted in a significant Line-type × Time interaction (Table 5) that was due to a reduction of the nose pokes in the H-BMR mice, as only in this line type the number of nose pokes significantly dropped once the bottle with tap water was replaced with the bottle containing quinine solution ($F_{1,308} = 5.1$; $p = 0.02$ for the planned comparison of the slope of change of the

number of nose pokes in the H-BMR mice vs. other line types; (Figure 5B). Thus, the H-BMR mice learned to avoid aversive cues faster than the mice of other line types.

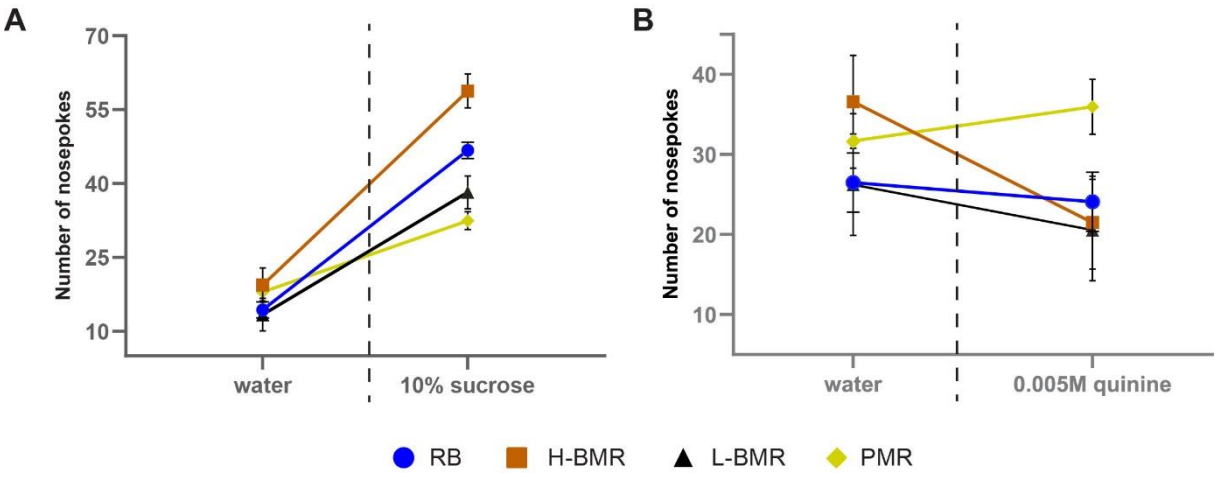

Figure 5. Results of the experiments in the IC system.

(A) Results of the reward-seeking discrimination learning: number of nose poke responses giving access to the bottle that contained tap water in preference learning, then sweetened water. Values are least square means (±SE) of nose pokes from the repeated measures mixed ANCOVA. When labelled with different letters, slopes at p = 0.05 (by a-priori contrasts). (B) The least square means (±SE) as in (A), but of the number of incorrect nose pokes counted in an aversive cue discrimination learning task, in which I used water solution of 0.005 M quinine.

Table 5. Repeated measures ANCOVA results for behavioral tests

	Line Type	**Day**	**Period**	**Day×Period**	**Line Type×Period**	**Line Type×Day**
Correct Nose pokes sucrose[a,b]	$F_{3,6} = 5.45$ $p = 0.03$	$F_{1,757}=345.5$ $p < 0.001$	$F_{1,757}=88.5$ $p < 0.001$	$F_{1, 757} = 89.2$ $p < 0.001$	$F_{1,} = 1.1$ $p = 0.3$	$F_{3, 757} = 19.8$ $p < 0.001$
Activity [b]	$F_{3,6} = 0.92$ $p = 0.48$	$F_{1,803}=21.3$ $p < 0.001$	$F_{1,803}=467.3$ $p < 0.001$	$F_{2,803}= 43.7$ $p < 0.001$	$F_{1,803} = 28.55$ $p < 0.001$	$F_{3,803} = 5.1$ $p = 0.002$
Licks[b,c]	$F_{3,6} = 0.36$ $p = 0. 78$	$F_{1,713}=204.3$ $p < 0.001$	$F_{1, 713}=259.0$ $p < 0.001$	$F_{2, 713} = 32.3$ $p < 0.001$	$F_{1, 713} = 5.3$ $p =0.001$	$F_{3, 713} = 5.0$ $p =0.002$
Incorrect Nose pokes quinine[a]	$F_{3,6} = 1.27$ $p = 0.36$	$F_{1,308} = 8.4$ $p = 0.004$	$F_{1,308} = 11.4$ $p <0.001$	$F_{1,308} = 4.1$ $p = 0.04$	$F_{1,308} = 2.35$ $p = 0.07$	$F_{3,308} = 2.88$ $p = 0.04$

[a] *the numbers of responses (i.e. nose pokes to the bottle with sucrose or quinine) were corrected for numbers of nose pokes to the bottle with tap water located in the same corner (used as a covariate, significant at $p < 0.001$).*

[b] *in this analysis the effect of a batch of animals simultaneously subjected to behavioral test was significant as a fixed factor ($p < 0.01$).*

[c] *the numbers of licks of the bottle with sucrose were corrected for numbers of licks to the bottle with tap water located in the same corner (used as a covariate, significant at $p < 0.001$)*

Finally, I investigated the differences in learning abilities among line types using a classic paradigm of contextual fear conditioning (Rustay, Browman, and Curzon 2008). Following conditioning elicited by a mild electric foot shock applied in a novel context I measured extinction of freezing response to a perceived threat in yet another group of naive mice. The Line type × Time interaction was statistically significant ($F_{15,30}$ = 2.27; p = 0.03), which reflected the heterogeneity of the dynamics of fear extinction in the studied line types, and accounted for the L-BMR mice losing fear response much faster than other line types (Figure 6).

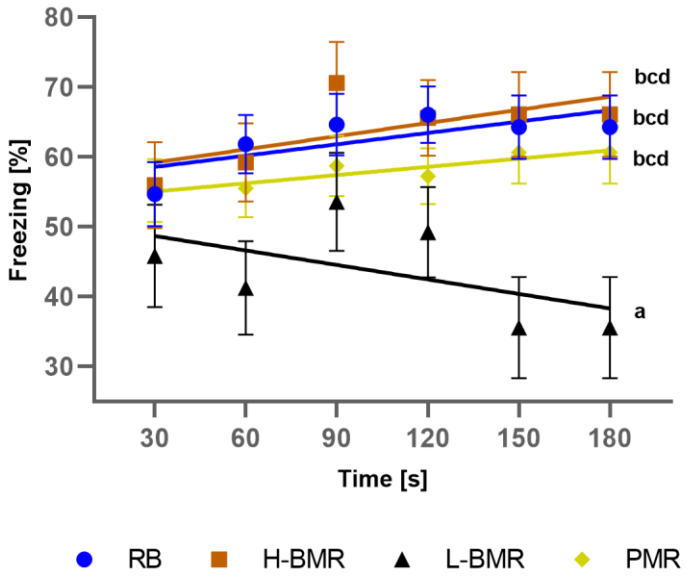

Figure 6. Changes in the fear response. Freezing (immobility) was measured as percentage [%] of total observations in 30 s intervals in high (H-BMR), low (L-BMR) Basal Metabolic Rate (BMR), Peak Metabolic Rate (PMR aka VO_2max) and random-bred (RB) line types of mice. Values are the least square means (±SE) from the repeated measures mixed ANCOVA. When labelled with different letters, slopes differed from at p = 0.05 (by a-priori contrast).

6.3 Discussion

According to the "expensive brain" hypothesis, the costs of increased brain size and cognitive abilities (CA) can be satisfied by (1) reallocation of resources towards brain growth and maintenance from other energetically expensive organs, as proposed by Aiello and Wheeler (1995); or by (2) increasing total energy intake, which may allow covering the costs of cognitive abilities without the need for reduction of other structures and functions, including digestive abilities (Isler and van Schaik 2006; Pontzer et al. 2016). Here, I comprehensively tested (1) and (2) in a mammalian experimental evolution model for the first time. My results do not support the existence of the brain-gut trade-off proposed in (1).

It is important to note, however, that throughout my experiment, mice were fed the same diet. Thus, compensation of the reduced gut by increased food quality could not be tested (Simmen et al. 2017). Nevertheless, at least in non-mammalian animal models, the brain-gut trade-off is likely to occur even without a shift in the quality of consumed food (Kotrschal et al. 2013). Also, as was demonstrated elsewhere (Książek, Czerniecki, and Konarzewski 2009) H-BMR mice have a considerable digestive safety margins, which would have allowed them for gut size reduction as suggested by the brain-gut trade-off hypothesis.

An increase in energy intake is the hallmark of the evolution of endothermy (Polymeropoulos, Oelkrug, and Jastroch 2018), mainly linked with the need to fuel reproduction (Koteja 2000). The H-BMR mice are characterized by increased energy intake and reproductive allocation (Sadowska et al., 2013) and increased gut mass (Table 1). These observations point to the second hypothesis and suggest that the selection for enhanced CA does not need to involve a brain-gut trade-off as an initial step toward the evolution of enhanced CA. On the contrary, the H-BMR mice with larger guts, but not brains, performed better in cognitively demanding tasks than their L-BMR counterparts and mice selected for maximum aerobic metabolism (PMR line type). Thus, the results of behavioral tests point to the positive association of CA with the evolution of BMR, rather than maximum aerobic metabolism (VO_{2max}, selected for in the PMR line type), which has also been implicated in the evolution of homeothermy and large brain size (A. F. Bennett and Ruben 1979; Lovegrove 2017; Raichlen and Gordon 2011)

Robustness of my findings critically depends on the evolutionary relevance of my animal model. To date studies on the evolution of brain size and function predominantly used comparative method which can only provide correlative inference, often confounded by phylogenetic effects. From a methodological perspective, a strong test of the genetic correlations between metabolic rates, CA, brain size and function should be provided by

artificial selection experiments, because they allow for manipulation of allele frequencies directly related to the expected associations (Garland and Rose 2009). Most importantly, animal models obtained through artificial selection experiments are characterized by diversification of the focal traits much exceeding the scope of variation characteristic for non-selected counterparts. Such diversification provides a robust substrate for statistical analyses making the statistical resolution of the analyzed associations possible. Thus, the line types originated from the same stock of Swiss Webster mice enables direct comparison of selection results, including the use of the RB line type as the common reference.

Apart from the biological relevance of the animal model used in my study, their robustness hinges on the credibility of the applied behavioral tests. To deliver biologically pertinent results, studies on CA in rodents must be carried out under conditions providing social and environmental stimulation, which triggers significant changes at cellular, molecular and behavioral levels, particularly in the cortex and hippocampus (Leggio et al. 2005). Here, I compared the behavior of mice of all line types when a highly rewarding 10% sucrose solution, or unpleasant taste of 0.005 M quinine solution appeared in the IC system. The H-BMR animals increased the number of nose pokes giving access to the sucrose solution and decreased nose pokes leading to the unpleasant bitter taste to a higher degree than the mice of other line types. The IC system allows for studying undisrupted behavior of mice in the familiar environment that exploits their natural preferences. Most of the conventional behavioral tests assess learning and memory by observation of a single animal in conditions that significantly differ from those in the home cage. Since in the IC housing and testing take place in the same cage, a familiar environment creates a unique opportunity to test voluntary behaviors of mice in low-stress conditions. Such conditions allow for testing learning and memory eliminating many confounding factors, such as stress related to novel environment, contact with a human experimenter and social isolation (Puścian et al. 2014). The system has been inspired by observations of freely moving mice in their natural environment, particularly genetically manipulated animals living in large outdoor pens and subjected to constantly changing environmental and social conditions (Abbott 2007). As a result, I demonstrated that IC offers a reliable range of tests assessing superior CA of the BMR mice in discrimination learning tasks. Conversely, contextual fear conditioning test, in which animals learned the association between the novel cage (new context) and unpleasant foot shocks revealed that L-BMR mice performed worse than the other animals.

In the first part of my studies tested directionality of the evolutionary relationships between energy expenditures, brain, gut, and CA in a mammalian model of experimental evolution. My results point to an evolutionary scenario that would involve initial selection for increased overall energy intake, which would result in positive genetic correlations with increased gut size and BMR (Kozłowski, Konarzewski, and Czarnoleski 2020) This selection may have involved an initial increase in neuronal plasticity, but not necessary increased brain size. This scenario is likely, if the brains built of more plastic neurons could have been more effective cognitively than the ones built of more neurons of lower plasticity (Herculano-Houzel 2011). Such brains would be more effective enabling foraging on better quality food. Subsequently, other trade-offs may have occurred in some lineages, such as proto-human apes. The trade-offs, such as, e.g., gut reduction, might have allowed for further brain size increases (Pontzer et al. 2016).

The above scenario obviously invites questions on the neurophysiological and histological underpinnings of the reported, between line type differences in CA. In the following chapters I will present the results of the analyses of neuronal plasticity, cell architecture and neuronal metabolic activity of the hippocampus- the region of brain most pertinent to behavioral tests presented therein.

7 Evolution of neurons size and function of the hippocampus: insights from the artificial selection experiment.

7.1 Materials & Methods

7.1.1 Animals

Animals were treated following the ethical standards of the European Union and Polish regulations. All procedures were approved by the Local Ethical Committee on Testing Animals. I used 3-4 month-old female mice from the selection experiment carried out at the Faculty of Biology, University of Bialystok. Measurements of LTP and histological analyses were carried out on separate groups of animals, randomly drawn from a stock of animals not qualified for further selection. Each mouse was randomly selected from a separate family belonging to a particular line.

The numbers of animals used in particular trials are given in Table 6:

Table 6. The number of animals included in the statistical analyses (n) and generation of origin (F). N.M.- not measured.

Trait measured	**RB**		**H-BMR**		**L-BMR**	
	n	F	n	F	n	F
LTP	8	38, 39	11	52, 53	8	52, 53
Nissl staining	14	38, 39	3	52, 53	3	52, 53
DAPI staining	4	59	4	60	4	60
Anti-GFAP immunostaining	18	59	3	59	4	59
Dendritic spine density	N.M.	-	5	60	5	60
Cytochrome oxidase (COO) activity staining	12	59	5	59	5	59

7.1.2 LTP measurements

To gain insight into the potential neuronal mechanism underlying observed differences in learning, I used long-term potentiation (LTP). LTP is an increase in the efficacy of signal

transmission between neurons caused by the strengthening of synapses by recent activity patterns. LTP is considered one of the primary cellular mechanisms underlying learning and memory (Cooke 2006). I compared the effects of repeated high-frequency stimulation of Schaeffer collaterals, that make excitatory synapses onto pyramidal cells in the CA1 area of the hippocampus, the brain structure crucial for spatial memory formation (Bliss and Collingridge 1993; Voikar et al. 2010).

Naive animals were anesthetized with isoflurane and decapitated. The brains were instantly removed and placed in cold artificial cerebrospinal fluid ACSF (NaCl 117 mM, $MgSO_4$ 1.2 mM, KCl 4.7 mM, $CaCl_2$ 2.5 mM, $NaHCO_3$ 25 mM, NaH_2PO_4 1.2 mM, 10 mM glucose, bubbled with carbogen) and both hemispheres were cut into 400 μm coronal slices with a vibratome (LeicaVT1000S). Slices containing the hippocampus were placed in a recording interface chamber (Harvard Apparatus) to recover for at least 1.5 h before the recordings start. The slices were continuously perfused with carbonated CSF at 33°C. Field excitatory postsynaptic potentials (fEPSPs), i.e., temporary depolarizations of postsynaptic membrane potentials, were recorded using a glass pipette filled with 20 mMNaCl (impendence 1.0–3.0 MΩ) from the *stratum radiatum* in CA1 area of the hippocampus. To evoke fEPSP, Schafer collateral-commissural afferents were stimulated every 30 s (test pulses at 0.033 Hz, 0.1 ms) with bipolar metal electrodes (FHC, USA). The intensity of test stimuli was adjusted to obtain fEPSP with slopes of one-third of the maximal response. After at least 15 min. of stable baseline, LTP was induced tetanically (three trains of 100 Hz, 1 s stimulation, separated by 3 min). After the end of the tetanic stimulation, a test pulse was applied for at least 90 min. Recordings were amplified (EX4-400 Dagan Corporation, USA), digitized (POWER1401, CED, UK) and slopes of fEPSP analyzed on-line and off-line. For analysis of LTP, the response slopes were expressed as a percentage of the average response slopes during the baseline period before LTP induction.

7.1.3 Staining

7.1.3.1 Perfusion and tissue preparation

BMR, PMR, and RB mice were deeply anesthetized with sodium pentobarbital (133.3 mg/ml), pentobarbital (26.7 mg/ml), and morbital, perfused transcardially with PBS (4°C, 50 ml per animal) and subsequently with 4% PFA (4°C, 70 ml per animal). Brains were dissected, postfixed in 4% PFA (o/n, 4°C) and cryoprotected in 30% sucrose solution (4°C) for at least 3 days. 40 μm coronal sections containing the hippocampus (coordinates following the atlas of

Paxinos and Franklin, 2001) were cut with the use of a cryostat (-20°C) and kept in PBS (4°C) until further handling. The number of all animals used in staining experiments is given in Table 6.

7.1.3.2 Nissl staining

Sections from -1,46; -1,82; -2,3 and -2,7 (mm from bregma) coronal sections from mouse hippocampus (coordinates following the atlas of Paxinos and Franklin, 2001) were mounted on gelatin-coated slides and then dried, dehydrated, and stained according to standard Nissl methods: slides were washed in 0.1 % cresyl violet solution (cresyl violet 0.1g; distilled water 50 ml; acetate buffer 0.2 M pH 3.6 50 ml) for 3-10 min and then dehydrated through graded alcohols (70, 95, 100, 100 %) and xylenes. Next, the slides were cleared in xylol and coverslipped with DePeX (Serva).

7.1.3.3 Measurements of the hippocampus and neuronal density estimation

All measurements were performed using CellSens Dimension Desktop (Olympus Corporation, Japan). The microscope image was captured using a light microscope Leica DM1000 LED connected to a Leica ICC50 camera, at 400x and 40x magnification for neuron counting and hippocampal measurements, respectively.

The area occupied by the pyramidal layer of the hippocampus was measured on each cross-section (as the pyramidal hippocampal area). Neuronal density was estimated in the pyramidal hippocampal layer, in CA1, CA2, and CA3 regions of the right and left side of hippocampus. The measurements were carried out in four coronal sections of hippocampus (-1.46; -1.82; -2.3 and -2.7 mm from bregma, coordinates following the atlas of Paxinos and Franklin, 2001).

The neurons were counted inside the rectangle frame with 3000 μm^2 area (counting frame) located at random in the central part of each of the regions. To obtain unbiased estimate of the neuron number within the frame I followed guidelines formulated by Gundersen (Gundersen 1977). Neurons crossed by the bottom and right-hand side borders of the frame were taken into account. Conversely, all neurons crossed by the left-hand side and upper borders of the frame were not accounted for thus, eight fields were analyzed for each hippocampal region.

7.1.3.4 DAPI staining and neuronal density estimation

Coronal sections of mouse hippocampus, positioned -2,06 mm from bregma (coordinates according to the atlas of Paxinos and Franklin, 2001), were mounted on gelatin-coated slides covered with Vectashield mounting medium containing DAPI. Next, the sections were photographed by means of a Zeiss LSM 780 confocal microscope, using a 63x lens for high

resolution imaging (Plan Apochromat 63x/1.4 Oil DIC, Zeiss) at identical capturing parameters for all images. Neuronal density was estimated in the pyramidal hippocampal layer, in CA1, and CA3 regions of the right and left side of hippocampus. The neurons were counted inside the rectangle frame with 2890 μm^2 area (counting frame) located at random in the central part of each of the regions. To obtain unbiased estimate of the neuron number within the frame I followed guidelines formulated by Gundersen (Gundersen 1977). Neurons crossed by the bottom and right-hand side borders of the frame were taken into account. Conversely, all neurons crossed by the left-hand side and upper borders of the frame were not accounted for thus, eight fields were analyzed for each hippocampal region. Number of neuronal nucleus in the CA1 and CA3 hippocampus region were counted by ImageJ (NIH) software.

7.1.3.5 Anti-GFAP immunostaining

Sections from the perfused brain were blocked with NDS (normal donkey serum). After the wash, sections were incubated in 0.1% Triton X-100 in PBS (PBST) for 15 minutes. Then, after preincubation in NDS-buffer (5% NDS in PBST) for 1 hour, the chicken anti-GFAP (Abcam, 1:800, Cambridge, MA, USA,) antibody (Ab) was added for overnight incubation at 4°C. After washing cycles in PBST (3 x 1 min), sections were incubated for 2 hours at room temperature with secondary antibody 647 Alexa Fluor (Jackson ImmunoResearch, West Grove, PA) at 1:400 concentration to detect GFAP. After the last round of washing with PBST (3 x 5 min, RT) slices were mounted (Fluoromount-G, Southern Biotech) on polysine slides. To control for non-specific binding and imaging artifacts ‘no primary antibody’ and ‘no secondary antibody’ immunostainings were carried out. Imaging was performed using Olympus microscope (Olympus Corporation, Japan) with RGB Camera Hamamatsu ORCA Flash 4.0 V2 and at 20x magnification for astrocyte counting. GFAP (astrocyte) immunoreactivity in the hippocampus (ROI) (in CA1-CA3 and DG regions) was assessed according to the ImageJ (NIH) macro presented in Supplementary materials.

7.1.3.6 Cytochrome oxidase (CCO) activity staining

Histochemical reaction for cytochrome oxidase was prepared according to a published protocol (Wong-Riley, 1979). Sections from -2,3 mm coronal section from mouse hippocampus (coordinates following the atlas of Paxinos and Franklin, 2001) were incubated in solution of 0.05 PB, 1 g sucrose, 50 mg ammonium nickel, 25 mg DAB, 15 mg cytochrome C, 10 mg catalase and 250 µl imidazole (amount per 100 ml). Slices were incubated for 6-8 h in 37°C and then washed (3 x 1 min) in 0.05 PB.

Cytochrome oxidase (CCO) activity in the CA1-CA3 and DG hippocampus region was assessed according to listed ImageJ (NIH) macro presented in Supplementary materials.

7.1.4 Dendritic spine analysis

Dendritic spines were visualized using the lipophilic dye Dil (1,1′-dioctadecyl-3,3,3′,3′-tetramethylindocarbocyanine perchlorate, #D282 Life Technologies, Warsaw, Poland). H- and L-BMR mice were deeply anesthetized with sodium pentobarbital (133,3 mg/ml), pentobarbital (26,7 mg/ml) and morbital. Their brains were dissected and fixed in 1.5% PFA (1h, RT) and moved to ice-cold PBS (10 min, 4°C). Next, they were cut into 140 μm sections on vibratome (Leica VT 1000S, Leica Biosystems Nussloch GmbH, Wetzlar, Germany). Slices were processed for Dil staining. Random dendrite labeling was performed using 1.6 μm tungsten particles (Bio-Rad, Hercules, CA, USA) coated with Dil. Dye was delivered to cells using the Gene Gun (Bio-Rad). After staining, slices were fixed with 0.4 % paraformaldehyde in phosphate-buffered saline (PBS; overnight at 4°C) and placed on microscopic slides. Z-stacks of dendrites from the CA1, CA3 and the DG regions of the hippocampus were acquired using the Zeiss LSM 880 confocal microscope with AiryScan on superresolution mode using a 63x objective for high resolution imaging (Plan Apochromat 63x/1.4 Oil DIC) (Zeiss, Poznań, Poland). Dil emission was excited using a HeNe 594 nm laser. For each image, the following parameters were applied: 70 nm pixel size, 300 nm Z-intervals, averaging 4. Maximum intensity projections of Z-stacks covering the dendrite length were analyzed using semiautomatic SpineMagick! software (Ruszczycki et al. 2012). It allows for marking the dendritic spine head and base manually. Next, the software automatically marks spine edges that can be adjusted manually to fully reflect the spine shape (Ruszczycki et al. 2012). In these experiments, the group size was 5 animals. For each animal, 7-10 single dendrites from selected brain areas (one dendrite per neuron per image) were analyzed and dendritic spine density was calculated. There were no clearly stained dendrites in the CA2 region, the preparation procedure was time consuming, so only BMR animals were taken into consideration.

7.1.5 Statistical analysis

Due to the technical complexity and labor-intensity of the LTP and spine density measurements I could not carry them out on animals from all line lines. In the case of the spine density, I restricted the comparisons to the H-BMR vs L-BMR line types. In case of LTP, however, I compared between the mice from the H-BMR, L-BMR line types, and animals randomly selected from one of the replicate random bred lines (belonging to the RB line type) as the reference. A comparison of body mass-corrected BMR of mice from this line with the animals

from the H-BMR and L-BMR lines revealed that BMR of the reference group fell between BMRs of the selected lines (ANCOVA, the effect of line: $F_{2,155} = 337.9$, $p < 0.001$). Body mass-corrected BMRs of mice under comparison averaged (LSM): 64.6 ± 0.7, 52.8 ± 1.8, 39.6 ± 0.7 for the H-BMR, reference group and L-BMR mice, respectively. Thus, this pattern of differences matches that found in a complete comparison of BMR presented in Figure 2A.

Data on LTP were arcsine transformed and then analyzed using repeated measures ANOVA with line type affiliation as a main factor. In this analysis, I compared the LTP slopes between the H-BMR and L-BMR line types along with one randomly drowned RB line as the reference group.

Data on spine density were analyzed using ANOVA with line type affiliation (H-BMR or L-BMR line type) and the hippocampal region as fixed factors and the respective interaction. Data on the remaining parameters were analyzed with the same ANOVA model but with replicated lines nested within line types as the random factor of the model (4 replications in the RB, but 1 line for H-BMR and L-BMR line types, respectively, as they were not replicated; 6 lines in total). The respective error mean square for 6 lines was used as the denominator of the F statistics testing the effect of line affiliation. Hence, the *df* for the between line type comparisons was 2 (for the F numerator) and 3 (for the denominator). Likewise, the *df* for pairwise t-test comparisons between the line types was 3.

In all initial analyses I also included lateralization (coded as a fixed factor), which was never significant ($p > 0.05$), and therefore was dropped from the final analyses.

7.1.5.1 Evaluation of the effect of genetic drift on LTP

As indicated above (Section 7.1.6), the data set on the LTP consisted of the measurements carried out on mice from the L-BMR, and H-BMR mice, along with the animals randomly selected from only one of the replicate random bred lines. Thus, unlike other comparisons between the L-BMR, and H-BMR mice vs. four randomly bred lines, this data set does not allow for an unequivocal exclusion of the effect of genetic drift as a potential source of the observed between line differences in LTP. To remedy this problem I compared the differences in LTP between the L-BMR and H-BMR mice using the guidelines recommended by Henderson (Henderson 1997) and Konarzewski et al. (Konarzewski 2005) presented in Section 6.1.5.3. Accordingly, I calculated the value of the d_{drift} for LTP and then compared it graphically with the values of d_{drift} calculated for different values of h^2. In calculating the d_{drift} I used the number of families equal to 80, and the coefficient of inbreeding $F = 30$.

7.2 Results

7.2.1 Neuronal plasticity measured with long-term potentiation (LTP)

In line with the behavioral results (presented in Section 6.2.2), the H-BMR mice manifested significantly higher neuronal plasticity measured as the LTP than the mice from the other two line types ($F_{2,24} = 18.4$ $p < 0.001$; Figure 7, Table 7). Furthermore, the difference in LTP between the H-BMR and L-BMR mice was more significant than that expected under genetic drift (Figure 8), suggesting a positive genetic correlation between BMR and LTP.

Table 7. Descriptive statistics of the LTP measurements (arithmetic means and standard deviations).

Trait measured	**Selection direction**							
	RB		**H-BMR**		**L-BMR**		**PMR**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
LTP (% of baseline)	191.7	29.5	241.9	45.0	151.0	28.8	N.M.	N.M

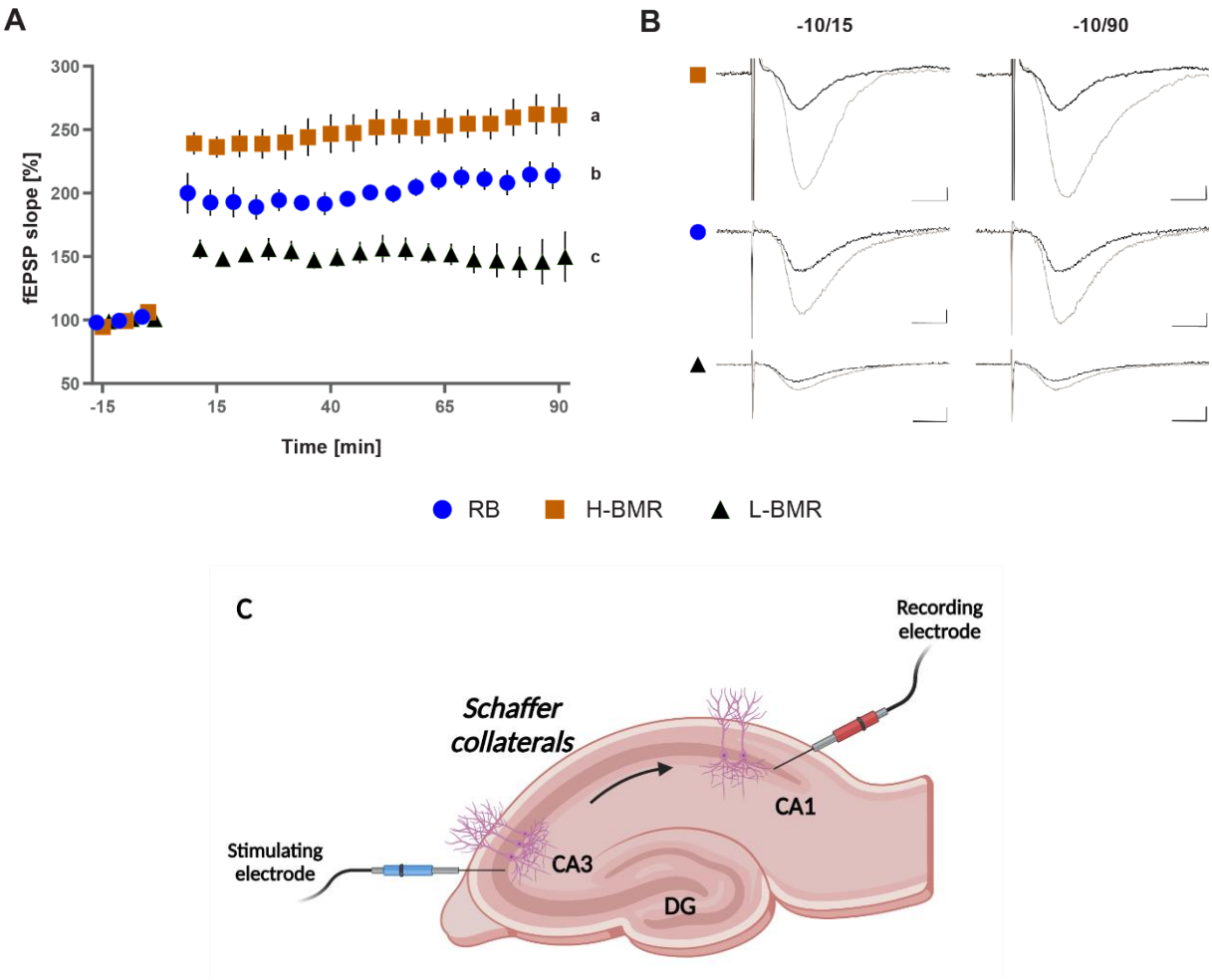

Figure 7. LTP recorded at the Shaffer collaterals of the hippocampus.

(A) The time course of maximal EPSP slopes was normalized to baseline in the CA1 region of the hippocampus. Long-term potentiation was induced by high-frequency stimulation (HFS; 3× 100 Hz) of the Schaffer collaterals in slices from the H-BMR (orange squares, n = 10), L-BMR (black triangle, n = 8), and the RB mouse line type (blue circles, n = 7). The slopes labeled with different letters differed at p = 0.05 (the Tukey post-hoc test). (B) Representative traces of fEPSP 10 min before (black) and 15 and 90 min after (grey) the induction of LTP. Scale bars = 2 mV and 5 ms. Values are the least square means (±SE) from the repeated measures mixed ANOVA. (C) Schematic diagram of Schaffer collateral projections from CA3 to CA1 area of mice hippocampus.

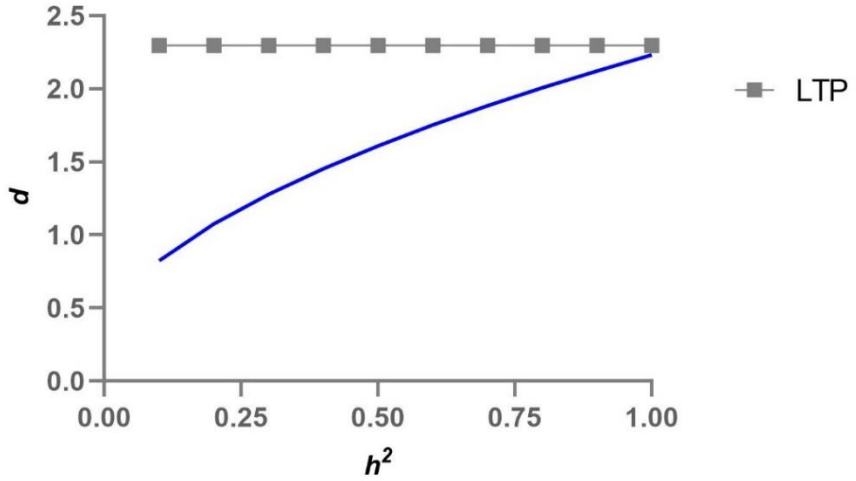

Figure 8. The magnitude of separation (d) of the H-BMR and L-BMR line types with respect to LTP, plotted against the upper boundary of the 95% confidence interval (d_{drift} , blue line). In the absence of available h^2 estimates, the value of d was plotted against the whole possible range of h^2.

7.2.2 Measurements of the hippocampus size and neuronal density of pyramidal layer

Total cross-sectional surface area of the hippocampus (being a proxy of its size) did not differ between the line types ($F_{2,3} = 0.01$; $p = 0.98$). Conversely, the line type effect on the pyramidal hippocampal area was significant ($F_{2,3} = 10.52$; $p = 0.04$), with the H-BMR mice having larger area than the RB and L-BMR animals ($p = 0.05$, t-test with df = 3 and the Tukey-Kramer correction, Figure 9). The remaining pairwise comparisons were not significant.

Neuronal density of the pyramidal layer (estimated by means of the Nissl staining) differed significantly between the line types (Figure 10, Table 8). Because of significant Line type × Region interaction (Table 8), I subsequently carried the between-line type comparisons separately within each of the CA1-CA3 regions of the layer. They revealed that the neuronal density of the H-BMR line type was higher than that of the RB line type in all of the regions (pairwise t-test with the Tukey-Kramer correction), whereas the L-BMR mice differed significantly from the RB line type only in the CA3 region (Figure 10). Neuronal densities of the mice of the L-BMR and H-BMR line types did not differ statistically in none of the pairwise comparisons.

To double check the robustness of estimates of neuronal density of the pyramidal layer obtained by means of the Nissl staining, I also carried out estimates of the density of the neurons stained with the DAPI method in the CA1 and CA3 regions. As in the case of the Nissl staining, the analysis revealed a significant Line type × Region interaction (Table 8). Separate analyses carried out within each of the regions did not detect the between line type differences in the CA1 region (Figure 11). Thus, unlike the Nissl staining, the DAPI staining did not reveal increased neuronal density in this region of hippocampus of H-BMR animals (compare Figure 10 and Figure 11). Conversely, both staining techniques yielded qualitatively identical results in the case of the CA3 region, with both the H-BMR and L-BMR mice having higher neuronal density than that found in the RB line type, and no significant difference between the H-BMR and L-BMR line types.

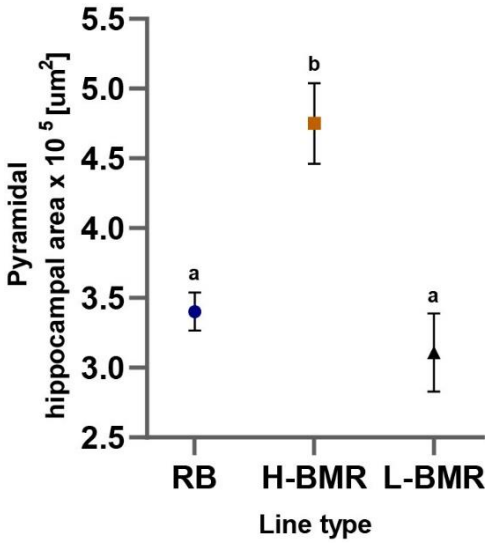

Figure 9. Measurement of pyramidal hippocampal area obtained by the Nissl staining. When labelled with different letters, means differed from each other at p = 0.05. Slices from the H-BMR (orange squares, n = 3), L-BMR (black triangle, n = 3), and the RB mouse line type (blue circles, n = 14).

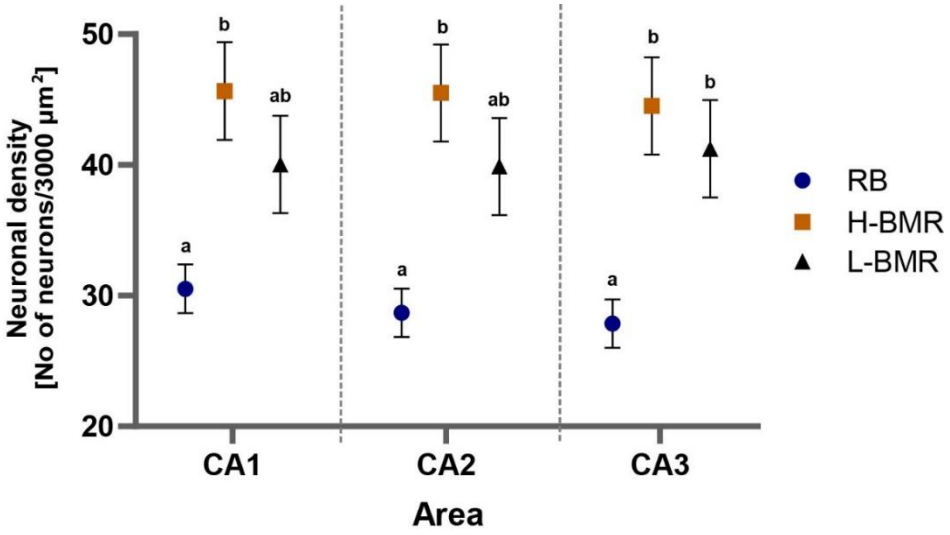

Figure 10. Neuronal density in the CA1, CA2 and CA3 regions of the pyramidal layer of hippocampus obtained by means of the Nissl staining. When labelled with different letters, means differed from each other at p = 0.05. Slices from the H-BMR (orange squares, n = 3), L-BMR (black triangle, n = 3), and the RB mouse line type (blue circles, n = 14).

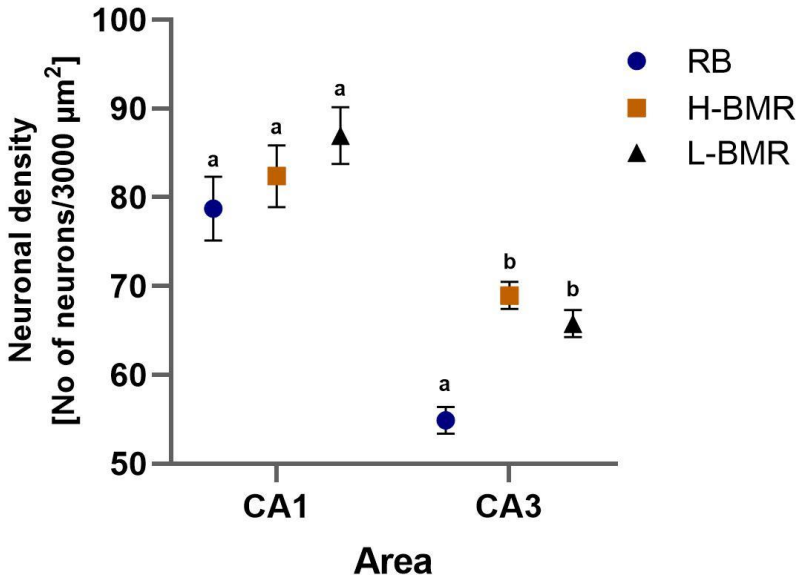

Figure 11. Neuronal density in the CA1 and CA3 regions of the pyramidal layer of hippocampus obtained by means of the DAPI staining. When labelled with different letters, means differed from each other at p = 0.05. Slices from the H-BMR (orange squares, n = 4), L-BMR (black triangle, n = 4), and the RB mouse line type (blue circles, n = 4).

7.2.3 Hippocampal astrocyte/neuron ratio

Differences between the line types in hippocampal astrocyte/neuron ratio (estimated as the percent of the GFAP signal positive cross- section area of hippocampus) were on the verge of statistical significance (Table 8). This result is, however, difficult to interpret because of the highly significant Line type × Region interaction (Table 8). Separate ANOVAs carried out within the Line type effect revealed that in the L-BMR and H-BMR line types, the CA3 region had a lower astrocyte/neuron ratio than in the CA1 and CA3 regions, respectively, with the remaining between-line type differences being non-significant at p = 0.05. The analyses carried out within the Region effect indicated that the only difference reaching statistical significance was that, within the CA2 region, in which the astrocyte/neuron ratio in the H-BMR line type was lower than that in the RB mice (Figure 12). It is, however worth noticing that the trend of reduced astrocyte/neuron ratio in the H-BMR and L-BMR vs. RB line type was apparent across all regions (Figure 12).

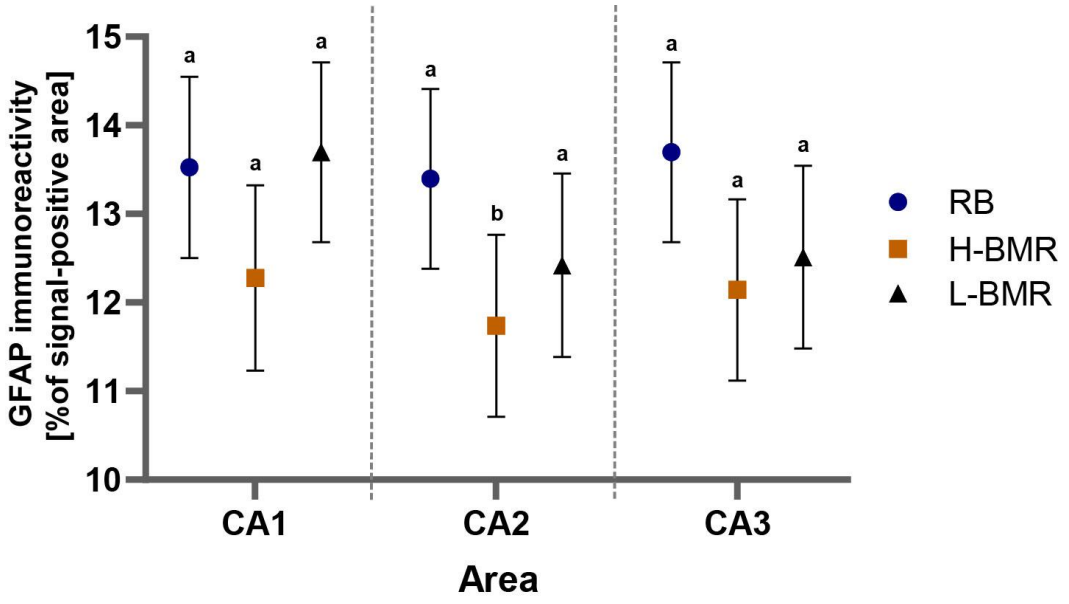

Figure 12. Astrocyte to neuron ratio, expressed as the percent of the GFAP signal positive cross- section CA1-CA3 areas of hippocampus. When labelled with different letters, means differed at p = 0.05. Slices from the H-BMR (orange squares, n = 3), L-BMR (black triangle, n = 4), and the RB mouse line type (blue circles, n = 18).

7.2.4 Dendritic spine density

As indicated by highly significant interaction (Table 8) dendritic spine density was significantly affected by both the hippocampal region and line type affiliation. Separate analyses carried out within regions revealed that the density did not differ between the line types in the CA1 and CA3 regions ($F_{1,88} = 0.0$ $p > 0.96$ and $F_{1,60} = 2{,}32$ $p > 013$, respectively). Conversely, the density was significantly higher in the DG region of the H-BMR mice, compared to the L-BMR line type ($F_{1,59} = 0.0$, $p = 0.0006$; Figure 13).

The difference in dendritic spine density in the DG between the H-BMR and L-BMR mice was more significant than that expected under genetic drift (Figure 14).

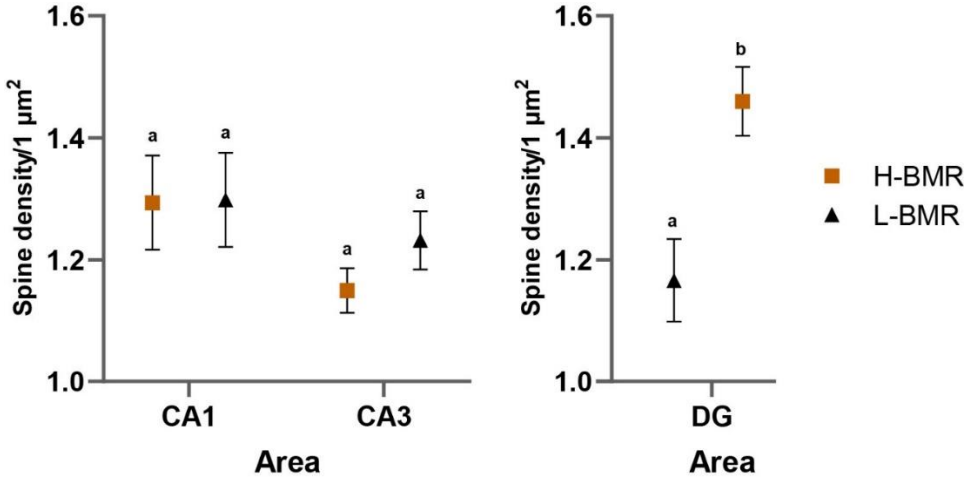

Figure 13. Dendritic spine density in the CA1, CA3 and DG region of hippocampus. When labelled with different letters, means differed at p = 0.05. Slices from the H-BMR (orange squares, n = 5), L-BMR (black triangle, n = 5).

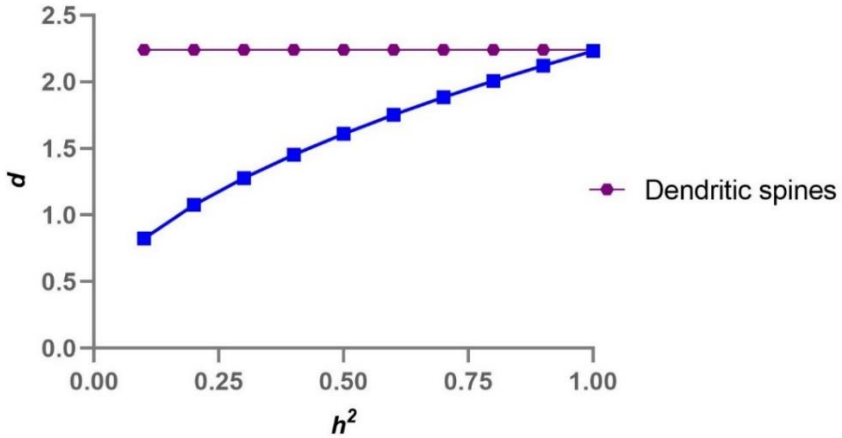

Figure 14. The magnitude of separation (d) of the H-BMR and L-BMR line types with respect to the density of dendritic spines in DG region of hippocampus. Value of d is positioned against the upper boundary of the 95% confidence interval (d_{drift} , blue line). In the absence of available h^2 estimates, the value of d was plotted against the whole possible range of h^2.

7.2.5 Cytochrome C oxidase (CCO) activity

Cytochrome C oxidase (CCO) density (a proxy of mitochondrial density) significantly varied between the line types and regions, with non-significant Line type × Region interaction (Table 8). The line type effect was clearly due to CCO density being much higher in the H-BMR line type than in other mice (Figure 15). A similar trend was also apparent in the DG region, though the difference did not reach statistical significance (Figure 15, Table 8)

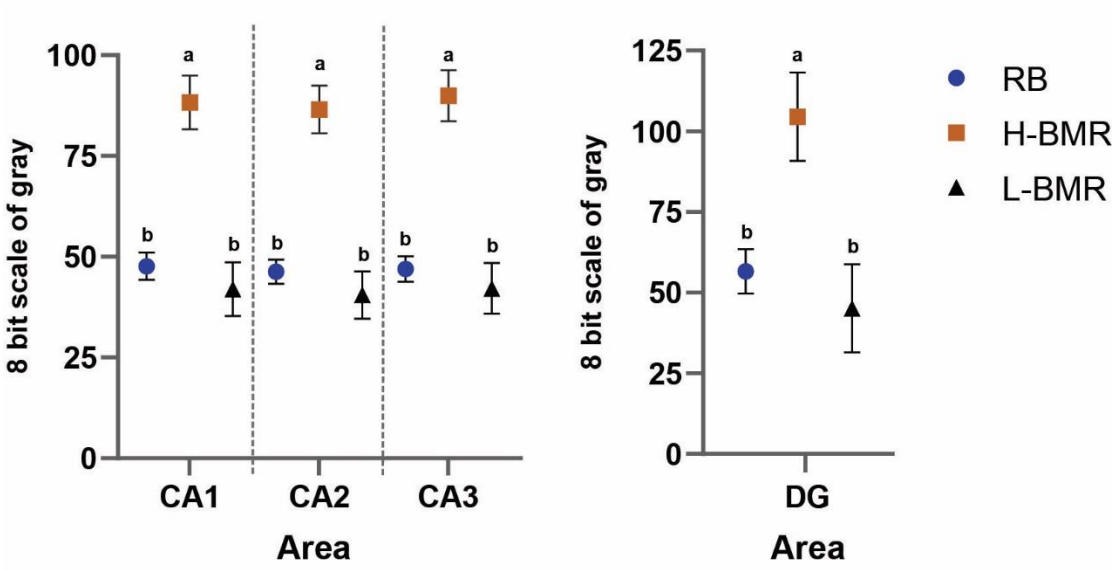

Figure 15. Cytochrome C oxidase density (expressed as the intensity of color in 8 bit scale of gray) in the CA1, CA2, CA3 and DG regions of hippocampus. When labelled with different letters, values of means differed at p = 0.05. Slices from the H-BMR (orange squares, n = 5), L-BMR (black triangle, n = 5), and the RB mouse line type (blue circles, n = 12).

Table 8. ANOVA results for the hippocampus anatomical and cellular characteristics. In all comparisons, the effect of lateralization was insignificant and thus dropped from the final models.

	Line type	**Region**	**Line type × Region**
Neuronal Density[a]	$F_{2,3} = 9.89$ $p = 0.048$	$F_{2,102} = 1.24$ $p = 0.30$	$F_{4,102} = 2.59$ $p = 0.04$
Neuronal Density[b]	$F_{2,7} = 6.94$ $p = 0.02$	$F_{1,109} = 228.4$ $p = 0.001$	$F_{2,109} = 6.29$ $p = 0.003$
Astrocyte/Neuron Surface ratio	$F_{2,3} = 9.40$ $p = 0.051$	$F_{2,111} = 5.52$ $p = 0.005$	$F_{4,111} = 3.90$ $p = 0.005$
Dendritic spine[c] density	$F_{1,222} = 2.44$ $p = 0.12$	$F_{2,222} = 3.01$ $p = 0.051$	$F_{2,222} = 9.49$ $p = 0.0001$
Cytochrome C oxidase in CA1-3	$F_{2,3} = 24.1$ $p = 0.01$	$F_{2,75} = 3.80$ $p = 0.03$	$F_{4,75} = 0.9$ $p = 0.1$
Cytochrome C oxidase in DG	$F_{2,3} = 6.42$ $p = 0.08$	-	-

[a] Nissl staining

[b] DAPI staining

[c] This comparison was carried out between the H-BMR and L-BMR line type

7.3 Discussion

Differences in CA among line types, demonstrated in behavioral tasks presented in Section 6.2.2 of my thesis, beg the question of the underlying neuronal mechanism. Both appetitive and aversive IC learning tasks involve the hippocampus (Voikar et al. 2010). Similarly, contextual fear conditioning relies on the hippocampus (Rudy and O'Reilly 2001). Since the H-BMR mice performed significantly better in the IC tests, while L-BMR mice were inferior in the fear conditioning test, I focused on the identification of the relevant neuronal mechanisms in the hippocampal neurophysiology, size, and cell architecture that could explain behavioral differences between those line types. Conversely, since the PMR line type did not stand out in any of the behavioral tests, I decided to drop them from further analysis, to avoid unjustified proliferation of degrees of freedom of statistical analyses.

My results revealed that the H-BMR, L-BMR, and RB mice differed in the studied parameters. In the LTP and spine density analyses the differences were very robust. However, morphological analyses brought results at the verge of the commonly adopted $p < 0.05$ level. It is, therefore, worth highlighting that the majority of my statistical inference is based on only 3 degrees of freedom, constituted by the respective line types, rather than the numbers of individual mice considered as degrees of freedom. Such an approach is the only proper way to avoid pseudoreplication and allow an appropriate distinction of the effects of selection from random effects, such as genetic drift (Henderson 1997). In the case of the LTP and spine density comparisons, which were not based on a complete data set incorporating all random bred lines (RB), I carried out additional analysis based on Henderson's guidelines (Henderson 1997). Thus, I am confident that our statistical inference is robust, despite the low levels of statistical significance of the reported tests.

First, I tested hippocampal neuronal plasticity using a well-established long-term potentiation (LTP) model (Bliss and Collingridge 1993). LTP is an increase in signal transmission between neurons caused by the strengthening of synapses by recent patterns of activity. LTP is considered one of the major cellular mechanisms underlying learning and memory formation (Cooke 2006). Furthermore, excitatory synaptic transmission requires ATP-dependent phosphorylation of AMPA receptors (Banke et al. 2000) and, therefore, should be positively associated with the rate of aerobic metabolism. My data show increased potentiation in the H-BMR mice and downregulation in the L-BMR animals, when compared to the RB mice. The results suggest upregulation of physiological processes intrinsic to learning in the H-BMR line

type, and their downregulation in L-BMR mice. It is consistent with the results of behavioral tests.

Second, I demonstrated that the neuronal density in the CA1-CA3 regions the pyramidal layer of hippocampus of the H-BMR mice was higher than that of the random bred (RB) animals that served as a reference. The area of pyramidal layer of the hippocampus was also larger in the H-BMR mice, which together with higher neuronal density translates into more numerous neurons. This finding is important, because higher number of neurons is generally associated with higher CA (Olkowicz et al. 2016).

In agreement with the above findings, dendritic spine density reflecting neuroplasticity was much higher in the dentate gyrus (DG) region of the H-BMR mice, compared to the L-BMR line type. Dendritic spines host synapses providing connectivity between neurons and undergo changes in their number and morphology driven by external stimuli (Runge, Cardoso, and de Chevigny 2020). The DG is the predominant point of entry of information reaching the hippocampus (Zhang, Schlögl, and Jonas 2020). Therefore the density of spines in this region is likely to affect learning and memory (Rebola, Carta, and Mulle 2017).

I must admit, however, that my results brought a couple of apparent conspicuous inconsistencies that are nevertheless potentially informative. First, although the H-BMR mice manifested significantly higher neuronal density than the RB (control) animals, pairwise comparisons indicated that its level did not statistically differ from the neuronal density found in the L-BMR mice. This results of the analysis of neuronal density obtained by means of the Nissl method were re-affirmed by an independent analysis with the use of DAPI staining. Thus, the trend apparent in results of neuronal density in the hippocampus suggests that divergent selection on BMR did not result in a parallel divergence in neuronal size. Importantly, this lack of divergence is paralleled by the lack of the between-line type difference in the cross-sectional area of the hippocampus, and therefore, its size. Thus, higher CA in the H-BMR mice than in the L-BMR animals is unlikely to be solely associated with numerous neurons. This conclusion agrees well with the results reported by Neves et al. (Neves et al. 2020). They did not observe significant correlations between the performance Swiss mice in behavioral tests and the number of neurons in any of their brain regions, including hippocampus. Thus, a broad scale positive correlation between the number of neurons and CA, reported by (Olkowicz et al. 2016), does not hold at the species level, at least in laboratory mice.

Another potential inconsistency of my results relates to much higher long-term potentiation (LTP) at the Schaffer collaterals of the H-BMR than the L-BMR mice, despite the lack of the between-line type difference in the dendritic spine density in the CA1 and CA3 regions. One could expect such a difference in those regions, since the spine formation and re-arrangement is considered critical for learning and memory formation (Lamprecht and LeDoux 2004). It is possible, however, that the lack of the difference in the CA1 and CA3 regions is compensated for by the higher spine density in the DG region of the H-BMR mice.

Overall, my results demonstrated that the differences in CA may arise through re-arrangement of the cell architecture and function in the hippocampus without a conspicuous differentiation of its size. The positive association between hippocampal size and CA has been demonstrated both in humans (Fotuhi, Do, and Jack 2012; Schinazi et al. 2013) and fish (Kotrschal et al. 2013). Therefore, a question arises, why the selection resulted in differentiation of CA and not scaling up of the hippocampus, and more generally, the whole brain, as observed interspecifically? This question is even more puzzling, when one considers that the evolutionary 'solution' of the re-arrangement, rather than up-sizing of the hippocampus, identified here in the H-BMR line type came at the substantial energetic cost, as indicated by high density of the cytochrome C oxidase (CCO) in the CA1-CA3 hippocampal regions.

To address above question I first need to discuss the legitimacy of the major assumption of my study, that is, the existence of a positive association between BMR and metabolic costs of maintenance of the brain tissue. Indeed, such association has been demonstrated in mammals at the interspecific level (Isler and van Schaik 2006). Moreover, in rodents, as much as 20% of interspecific variation in body mass-corrected BMR can be attributed to brain mass variation (Sobrero et al. 2011). Since BMR reflects the costs of maintenance of metabolically expensive organs (Konarzewski and Diamond 1995), it must also include neuronal energy consumption related to maintaining the resting potential and its restoration following action potential propagation (Niven and Laughlin 2008). As I pointed out in the Introduction, however, to be effective as an evolutionary mechanism driving encephalization, the BMR-brain size association must be present at the species level, where natural selection operates (Bennett & Harvey, 1985; Isler & van Schaik, 2006). Our results show that, at least in our animal model, the positive genetic correlation between BMR and brain size is absent. However, divergent selection for BMR resulted in a correlated response of an increased CA in the H-BMR mice. Thus, enhanced CA are indeed associated with high BMR at the species level, but not necessarily due to an increased brain size.

I cannot quantify the exact contribution of an increased CA of the H-BMR mice to their exceptional BMR. Yet, their high neuronal density in the CA1-CA3 regions may partly account for this association. Interestingly, the cytochrome C oxidase (CCO) density of the H-BMR mice is much higher than that of the L-BMR ones, despite the lack of the between-line type difference in neuronal density. The cytochrome C oxidase (CCO) can be considered a metabolic marker of neuronal activity (Wong-Riley, 2012). Together, these results strongly suggest that individual neurons of the H-BMR mice have a much higher metabolic rate than those of the L-BMR line type, which agrees well with their significantly higher LTP. This result also suggests that divergent selection for BMR indirectly increased the metabolic rate of individual neurons and that this increase, rather than an increase in the total number of neurons, may account for enhanced CA of the H-BMR mice. My results therefore call into question the constancy of energy expenditures of a single neuron inferred by Herculano-Houzel (Herculano-Houzel 2011) based on interspecific comparisons.

In conclusion, my findings demonstrated several associations that run counter interspecific patterns widely considered as mechanisms of the evolution of CA and large brains. My results, along with other studies carried out interspecifically e.g. (Benson-Amram et al., 2016) should send a warning signal across the neuroscience community, which is predominantly concentrated on interspecific comparisons as a source of inference (Deaner, Nunn, and Van Schaik 2000). Such inference can be cross-checked by the experimental evolution models like the one used here because, unlike comparative studies, they provide a rigorous and biologically meaningful means for identification of causal relationships.

8 Conclusions

- My study does not support the existence of the brain-gut trade-offs postulated by the "expensive tissue" hypothesis. The results of behavioral tests support the existence of the link between cognitive abilities (CA) and high basal metabolic rate (BMR) as postulated by the "expensive brain" concept. My study provide the first experimental test of both hypotheses carried out by means of a mammalian animal model;

- Selection on high basal metabolic rate (BMR) increases neuronal plasticity, which may be the first step towards increased cognitive abilities (CA), instead of the brain size increase through the number or size of neurons. Thus, my results run counter a general positive correlation between CA and brain size observed at the interspecific level. They therefore call into question the validity of those correlations as a universal pattern of the evolution of encephalization;

- My results demonstrated that the differences in cognitive abilities (CA) may arise through re-arrangement of the cell architecture and function of the hippocampus, without a conspicuous differentiation of its size. I therefore concur with still scant studies criticizing an overly simplified assumption of a strict size-function relationship frequently used in comparative studies on the evolution of brain size and CA;

- Divergent selection for BMR indirectly increased the metabolic rate of individual neurons. Thus, my results do not support the notion of the size- independence of energy budgets of neurons. An increase of their aerobic metabolism, along with an increase of the total number of smaller neurons, may account for an enhanced cognitive abilities CA of the H-BMR mice.

9 Bibliography

1. Abbott, Alison. 2007. “A Question of Survival: International Collaboration and a Can-Do Spirit Have Allowed Some Russian Scientists to Flourish. Alison Abbott Watches an Extraordinary Field Test for Mutant Mice in the Russian Wilderness.” *Nature* 449 (7162): 532–35. https://go.gale.com/ps/i.do?p=HRCA&sw=w&issn=00280836&v=2.1&it=r&id=GALE%7CA189796708&sid=googleScholar&linkaccess=fulltext.
2. Aiello, Leslie C., and Peter Wheeler. 1995. “The Expensive-Tissue Hypothesis: The Brain and the Digestive System in Human and Primate Evolution.” *Current Anthropology* 36 (2): 199–221. https://doi.org/10.1086/204350.
3. Allman, J., T. McLaughlin, and A. Hakeem. 1993. “Brain Weight and Life-Span in Primate Species.” *Proceedings of the National Academy of Sciences* 90 (1): 118–22. https://doi.org/10.1073/PNAS.90.1.118.
4. Armstrong, E., and R. Bergeron. 1985. “Relative Brain Size and Metabolism in Birds.” *Brain, Behavior and Evolution* 26 (3–4): 141–53. https://doi.org/10.1159/000118782.
5. Atchley, William R., Bruce Riska, Luci A. P. Kohn, A. Alison Plummer, and J. J. Rutledge. 1984. “A Quantitative Genetic Analysis of Brain and Body Size Associations, Their Origin and Ontogeny: Data From Mice.” *Evolution* 38 (6): 1165. https://doi.org/10.2307/2408625.
6. Attwell, David, and Costantino Iadecola. 2002. “The Neural Basis of Functional Brain Imaging Signals.” *Trends in Neurosciences* 25 (12): 621–25. https://doi.org/10.1016/S0166-2236(02)02264-6.
7. Attwell, David, and Simon B. Laughlin. 2001. “An Energy Budget for Signaling in the Grey Matter of the Brain.” *Journal of Cerebral Blood Flow and Metabolism* 21 (10): 1133–45. https://doi.org/10.1097/00004647-200110000-00001.
8. Banke, T. G., D. Bowie, H.-K. Lee, R. L. Huganir, A. Schousboe, and S. F. Traynelis. 2000. “Control of GluR1 AMPA Receptor Function by CAMP-Dependent Protein Kinase.” *The Journal of Neuroscience* 20 (1): 89–102. https://doi.org/10.1523/JNEUROSCI.20-01-00089.2000.
9. Bennett, Albert F., and John A. Ruben. 1979. “Endothermy and Activity in Vertebrates.” *Science* 206 (4419): 649–54. https://doi.org/10.1126/SCIENCE.493968.
10. Bennett, Peter M., and Paul H. Harvey. 1985. “Brain Size, Development and

Metabolism in Birds and Mammals." *Journal of Zoology* 207 (4): 491–509. https://doi.org/10.1111/J.1469-7998.1985.TB04946.X.

11. Benson-Amram, Sarah, Ben Dantzer, Gregory Stricker, Eli M. Swanson, and Kay E. Holekamp. 2016. "Brain Size Predicts Problem-Solving Ability in Mammalian Carnivores." *Proceedings of the National Academy of Sciences of the United States of America* 113 (9): 2532–37. https://doi.org/10.1073/PNAS.1505913113/SUPPL_FILE/PNAS.1505913113.SM01.MP4.
12. Biggar, Kyle K., and Kenneth B. Storey. 2015. "Insight into Post-Transcriptional Gene Regulation: Stress-Responsive MicroRNAs and Their Role in the Environmental Stress Survival of Tolerant Animals." *Journal of Experimental Biology* 218 (9): 1281–89. https://doi.org/10.1242/JEB.104828.
13. Bliss, T. V. P., and G. L. Collingridge. 1993. "A Synaptic Model of Memory: Long-Term Potentiation in the Hippocampus." *Nature* 361 (6407): 31–39. https://doi.org/10.1038/361031a0.
14. Burini, Roberto Carlos, and William R. Leonard. 2018. "The Evolutionary Roles of Nutrition Selection and Dietary Quality in the Human Brain Size and Encephalization." *Nutrire 2018 43:1* 43 (1): 1–9. https://doi.org/10.1186/S41110-018-0078-X.
15. Carlo, C. Nikoosh, and Charles F. Stevens. 2013. "Structural Uniformity of Neocortex, Revisited." *Proceedings of the National Academy of Sciences of the United States of America* 110 (4): 1488–93. https://doi.org/10.1073/PNAS.1221398110/SUPPL_FILE/PNAS.201221398SI.PDF.
16. Chambers, Helen Rebecca, Sandra Andrea Heldstab, and Sean J. O'Hara. 2021. "Why Big Brains? A Comparison of Models for Both Primate and Carnivore Brain Size Evolution." *PLoS ONE* 16 (12 December). https://doi.org/10.1371/journal.pone.0261185.
17. Chappell, Mark A., Theodore Garland, Geoff F. Robertson, and Wendy Saltzman. 2007. "Relationships among Running Performance, Aerobic Physiology and Organ Mass in Male Mongolian Gerbils." *Journal of Experimental Biology* 210 (23): 4179–97. https://doi.org/10.1242/JEB.006163.
18. Chrząścik, Katarzyna M., Edyta T. Sadowska, Agata Rudolf, and Paweł Koteja. 2014. "Learning Ability in Bank Voles Selected for High Aerobic Metabolism, Predatory Behaviour and Herbivorous Capability." *Physiology & Behavior* 135 (August): 143–

51. https://doi.org/10.1016/J.PHYSBEH.2014.06.007.

19. Cooke, S. F. 2006. “Plasticity in the Human Central Nervous System.” *Brain* 129 (7): 1659–73. https://doi.org/10.1093/brain/awl082.
20. Curz, Peter, Nathan R. Rustay, and Kaitlin E. Browman. 2009. “Chapter 2 Cued and Contextual Fear Conditioning f or Rodents.” *Methods of Behavior Analysis in Neuroscience*.
21. Deaner, Robert O., Karin Isler, Judith Burkart, and Carel Van Schaik. 2007. “Overall Brain Size, and Not Encephalization Quotient, Best Predicts Cognitive Ability across Non-Human Primates.” *Brain, Behavior and Evolution* 70 (2): 115–24. https://doi.org/10.1159/000102973.
22. Deaner, Robert O., Charles L. Nunn, and Carel P. Van Schaik. 2000. “Comparative Tests of Primate Cognition: Different Scaling Methods Produce Different Results.” *Brain, Behavior and Evolution* 55 (1): 44–52. https://doi.org/10.1159/000006641.
23. Dunbar, Robin I M. n.d. “The Social Brain Hypothesis.” https://doi.org/10.1002/(SICI)1520-6505(1998)6:5.
24. Eisenberg, John F., and Don E. Wilson. 1978. “Relative Brain Size and Feeding Strategies in the Chiroptera.” *Evolution* 32 (4): 740. https://doi.org/10.2307/2407489.
25. Erbslöh, F., A. Bernsmeier, and H. R. Hillesheim. 1958. “[The Glucose Consumption of the Brain & Its Dependence on the Liver].” *Archiv Fur Psychiatrie Und Nervenkrankheiten, Vereinigt Mit Zeitschrift Fur Die Gesamte Neurologie Und Psychiatrie* 196 (6): 611–26. https://doi.org/10.1007/BF00344388.
26. Fotuhi, Majid, David Do, and Clifford Jack. 2012. “Modifiable Factors That Alter the Size of the Hippocampus with Ageing.” *Nature Reviews Neurology 2012 8:4* 8 (4): 189–202. https://doi.org/10.1038/nrneurol.2012.27.
27. Garland, Theodore, and Michael R. Rose. 2009. *Experimental Evolution: Concepts, Methods, and Applications of Selection Experiments*. *Experimental Evolution: Concepts, Methods, and Applications of Selection Experiments*. https://doi.org/10.5860/choice.47-6844.
28. Gębczyński, Andrzej, and Marek Konarzewski. 2009. “Metabolic Correlates of Selection on Aerobic Capacity in Laboratory Mice: A Test of the Model for the Evolution of Endothermy.” *J. Exp. Biol.* 212: 2872–78. https://doi.org/10.1111/j.1420-9101.2009.01734.x.
29. Gray, E. G. 1959. “Electron Microscopy of Synaptic Contacts on Dendrite Spines of the Cerebral Cortex.” *Nature 1959 183:4675* 183 (4675): 1592–93.

https://doi.org/10.1038/1831592a0.

30. Gundersen, Hans Jørgen G. 1977. “Notes on the Estimation of the Numerical Density of Arbitrary Profiles: The Edge Effect.” *Journal of Microscopy* 111 (2): 219–23. https://doi.org/10.1111/J.1365-2818.1977.TB00062.X.
31. Harris, Julia J., Renaud Jolivet, and David Attwell. 2012. “Synaptic Energy Use and Supply.” *Neuron* 75 (5): 762–77. https://doi.org/10.1016/J.NEURON.2012.08.019.
32. Harvey, P. H., T. H. Clutton-Brock, and G. M. Mace. 1980. “Brain Size and Ecology in Small Mammals and Primates.” *Proceedings of the National Academy of Sciences* 77 (7): 4387–89. https://doi.org/10.1073/PNAS.77.7.4387.
33. Harvey, P H, and M D Pagel. 1992. “The Comparative Method in Evolutionary Biology.” *Journal of Classification 1992 9:1* 9 (1): 169–72. https://doi.org/10.1007/BF02618482.
34. Healy, Susan D., and Candy Rowe. 2013. “Costs and Benefits of Evolving a Larger Brain: Doubts over the Evidence That Large Brains Lead to Better Cognition.” *Animal Behaviour* 86 (4): e1. https://doi.org/10.1016/J.ANBEHAV.2013.05.017.
35. Healy, Susan D, and Candy Rowe. 2007. “A Critique of Comparative Studies of Brain Size.” *Proceedings of the Royal Society B: Biological Sciences* 274 (1609): 453–64. https://doi.org/10.1098/rspb.2006.3748.
36. Henderson, Norman D. 1997. “Spurious Associations in Unreplicated Selected Lines.” *Behavior Genetics* 27 (2): 145–54. https://doi.org/10.1023/A:1025689425738.
37. Herculano-Houzel, Suzana. 2011. “Scaling of Brain Metabolism with a Fixed Energy Budget per Neuron: Implications for Neuronal Activity, Plasticity and Evolution.” *PLoS ONE* 6 (3). https://doi.org/10.1371/JOURNAL.PONE.0017514/PONE_0017514_PDF.PDF.
38. Herculano-Houzel, Suzana, Kamilla Avelino-de-Souza, Kleber Neves, Jairo Porfírio, Débora Messeder, Larissa Mattos Feijó, José Maldonado, and Paul R. Manger. 2014. “The Elephant Brain in Numbers.” *Frontiers in Neuroanatomy* 8 (JUN): 46. https://doi.org/10.3389/FNANA.2014.00046/BIBTEX.
39. Hill, William G., and Trudy F.C. Mackay. 2004. “D. S. Falconer and Introduction to Quantitative Genetics.” *Genetics* 167 (4): 1529–36. https://doi.org/10.1093/GENETICS/167.4.1529.
40. Hofman, M. A. 2015. “Energy Metabolism, Brain Size and Longevity in Mammals.” *Https://Doi.Org/10.1086/413544* 58 (4): 495–512. https://doi.org/10.1086/413544.
41. Holtmaat, Anthony, and Karel Svoboda. 2009. “Experience-Dependent Structural

Synaptic Plasticity in the Mammalian Brain." *Nature Reviews Neuroscience 2009 10:9* 10 (9): 647–58. https://doi.org/10.1038/nrn2699.

42. Hooper, Rebecca, Becky Brett, and Alex Thornton. 2022. "Problems with Using Comparative Analyses of Avian Brain Size to Test Hypotheses of Cognitive Evolution." Edited by Vitor Hugo Rodrigues Paiva. *PloS One* 17 (7): e0270771. https://doi.org/10.1371/JOURNAL.PONE.0270771.
43. Isler, Karin, and Carel P. van Schaik. 2009. "The Expensive Brain: A Framework for Explaining Evolutionary Changes in Brain Size." *Journal of Human Evolution* 57 (4): 392–400. https://doi.org/10.1016/J.JHEVOL.2009.04.009.
44. Isler, Karin, and Carel P van Schaik. 2006. "Metabolic Costs of Brain Size Evolution." *Biology Letters* 2 (4): 557–60. https://doi.org/10.1098/rsbl.2006.0538.
45. Jerison, Harry J. 1973. *Evolution of the Brain and Intelligence*. Academic Press.
46. Jolly, Alison. 1966. "Lemur Social Behavior and Primate Intelligence." *Science* 153 (3735): 501–6. https://doi.org/10.1126/SCIENCE.153.3735.501.
47. Kandel, Eric R. 2012. "The Molecular Biology of Memory: CAMP, PKA, CRE, CREB-1, CREB-2, and CPEB." *Molecular Brain* 5 (1): 1–12. https://doi.org/10.1186/1756-6606-5-14/METRICS.
48. Karbowski, Jan. 2019. "Metabolic Constraints on Synaptic Learning and Memory." *Journal of Neurophysiology* 122 (4): 1473–90. https://doi.org/10.1152/JN.00092.2019/ASSET/IMAGES/LARGE/Z9K0091952010009.JPEG.
49. Kiryk, Anna, Gabriela Mochol, Robert K. Filipkowski, Marcin Wawrzyniak, Victoria Lioudyno, Ewelina Knapska, Tomasz Gorkiewicz, et al. 2011. "Cognitive Abilities of Alzheimers Disease Transgenic Mice Are Modulated by Social Context and Circadian Rhythm." *Current Alzheimer Research* 8 (8): 883–92. https://doi.org/10.2174/156720511798192745.
50. Knapska, E., V. Lioudyno, A. Kiryk, M. Mikosz, T. Gorkiewicz, P. Michaluk, M. Gawlak, et al. 2013. "Reward Learning Requires Activity of Matrix Metalloproteinase-9 in the Central Amygdala." *Journal of Neuroscience* 33 (36): 14591–600. https://doi.org/10.1523/JNEUROSCI.5239-12.2013.
51. Kolb, E. M., E. L. Rezende, L. Holness, A. Radtke, S. K. Lee, A. Obenaus, and T. Garland. 2013. "Mice Selectively Bred for High Voluntary Wheel Running Have Larger Midbrains: Support for the Mosaic Model of Brain Evolution." *The Journal of Experimental Biology* 216 (Pt 3): 515–23. https://doi.org/10.1242/JEB.076000.

52. Konarzewski, M. 2005. “Artificial Selection on Metabolic Rates and Related Traits in Rodents.” *Integrative and Comparative Biology* 45 (3): 416–25. https://doi.org/10.1093/icb/45.3.416.
53. Konarzewski, M., and J. Diamond. 1995. “Evolution of Basal Metabolic Rate and Organ Masses in Laboratory Mice.” *Evolution* 49 (6): 1239–48. https://doi.org/10.1111/j.1558-5646.1995.tb04450.x.
54. Koteja, Paweł. 2000. “Energy Assimilation, Parental Care and the Evolution of Endothermy.” *Proceedings of the Royal Society of London. Series B: Biological Sciences* 267 (1442): 479–84. https://doi.org/10.1098/RSPB.2000.1025.
55. Kotrschal, Alexander, Björn Rogell, Andreas Bundsen, Beatrice Svensson, Susanne Zajitschek, Ioana Brännström, Simone Immler, Alexei A. Maklakov, and Niclas Kolm. 2013. “Artificial Selection on Relative Brain Size in the Guppy Reveals Costs and Benefits of Evolving a Larger Brain.” *Current Biology*. https://doi.org/10.1016/j.cub.2012.11.058.
56. Kowalski, Jakub M, Alicja Puścian, Zofia Mijakowska, Maria Nalberczak, Kasia Radwańska, and Szymon Łęski. 2015. “PyMICE - a Python™ Library for Analysis of Mice Behaviour.” *BMC Neuroscience* 16 (S1): P145. https://doi.org/10.1186/1471-2202-16-S1-P145.
57. Kozłowski, Jan, Marek Konarzewski, and Marcin Czarnoleski. 2020. “Coevolution of Body Size and Metabolic Rate in Vertebrates: A Life-history Perspective.” *Biological Reviews* 95 (5): 1393–1417. https://doi.org/10.1111/brv.12615.
58. Książek, Aneta, Jan Czerniecki, and Marek Konarzewski. 2009. “Phenotypic Flexibility of Traits Related to Energy Acquisition in Mice Divergently Selected for Basal Metabolic Rate (BMR).” *Journal of Experimental Biology* 212 (6): 808–14. https://doi.org/10.1242/jeb.025528.
59. Książek, Aneta, Marek Konarzewski, and Iwona B. Łapo. 2004. “Anatomic and Energetic Correlates of Divergent Selection for Basal Metabolic Rate in Laboratory Mice.” *Physiological and Biochemical Zoology* 77 (6): 890–99. https://doi.org/10.1086/425190.
60. Lamprecht, Raphael, and Joseph LeDoux. 2004. “Structural Plasticity and Memory.” *Nature Reviews Neuroscience 2004 5:1* 5 (1): 45–54. https://doi.org/10.1038/nrn1301.
61. Leggio, Maria Giuseppa, Laura Mandolesi, Francesca Federico, Francesca Spirito, Benedetta Ricci, Francesca Gelfo, and Laura Petrosini. 2005. “Environmental Enrichment Promotes Improved Spatial Abilities and Enhanced Dendritic Growth in

the Rat." *Behavioural Brain Research* 163 (1): 78–90. https://doi.org/10.1016/J.BBR.2005.04.009.

62. Lewin, Roger. 1982. "How Did Humans Evolve Big Brains?" *Science* 216 (4548): 840–41. https://doi.org/10.1126/SCIENCE.7079741/ASSET/03B2C949-D928-4128-A9F7-CCC899933F73/ASSETS/SCIENCE.7079741.FP.PNG.
63. Liao, Wen Bo, Shang Ling Lou, Yu Zeng, and Alexander Kotrschal. 2016. "Large Brains, Small Guts: The Expensive Tissue Hypothesis Supported within Anurans." *Https://Doi.Org/10.1086/688894* 188 (6): 693–700. https://doi.org/10.1086/688894.
64. Lipp, Hans-peter. 2005. "High-Throughput and Automated Behavioural Screening of Normal and Genetically Modified Mice." *Changes* 4 (2000): 2000–2004.
65. Lipp, Hans-Peter, O Litvin, Mj Galsworthy, D L Vyssotski, a L Vyssotski, P Zinn, and a E Rau. 2005. "Automated Behavioral Analysis of Mice Using INTELLICAGE: Inter-Laboratory Comparisons and Validation with Exploratory Behavior and Spatial Learning." *Proceedings of Measuring Behaviour 2005* 2005 (September): 66–69.
66. Lovegrove, Barry G. 2017. "A Phenology of the Evolution of Endothermy in Birds and Mammals." *Biological Reviews* 92 (2): 1213–40. https://doi.org/10.1111/brv.12280.
67. Mace, Georgina M., Paul H. Harvey, and T. H. Clutton-Brock. 1981. "Brain Size and Ecology in Small Mammals." *Journal of Zoology* 193 (3): 333–54. https://doi.org/10.1111/J.1469-7998.1981.TB03449.X.
68. Magistretti, Pierre J. 2009. "Role of Glutamate in Neuron-Glia Metabolic Coupling." *The American Journal of Clinical Nutrition* 90 (3): 875S-880S. https://doi.org/10.3945/AJCN.2009.27462CC.
69. Magistretti, Pierre J., and Igor Allaman. 2015. "A Cellular Perspective on Brain Energy Metabolism and Functional Imaging." *Neuron* 86 (4): 883–901. https://doi.org/10.1016/J.NEURON.2015.03.035.
70. Magistretti, Pierre J., Luc Pellerin, Douglas L. Rothman, and Robert G. Shulman. 1999. "Energy on Demand." *Science* 283 (5401): 496–97. https://doi.org/10.1126/SCIENCE.283.5401.496.
71. Masuda, Akira, Yuki Kobayashi, and Shigeyoshi Itohara. 2018. "Automated, Long-Term Behavioral Assay for Cognitive Functions in Multiple Genetic Models of Alzheimer's Disease, Using IntelliCage." *Journal of Visualized Experiments*, no. 138 (August). https://doi.org/10.3791/58009.
72. McKenna, Mary C., Gerald A. Dienel, Ursula Sonnewald, Helle S. Waagepetersen,

and Arne Schousboe. 2012. “Energy Metabolism of the Brain.” *Basic Neurochemistry*, January, 200–231. https://doi.org/10.1016/B978-0-12-374947-5.00011-0.

73. Mergenthaler, Philipp, Ute Lindauer, Gerald A. Dienel, and Andreas Meisel. 2013. “Sugar for the Brain: The Role of Glucose in Physiological and Pathological Brain Function.” *Trends in Neurosciences* 36 (10): 587–97. https://doi.org/10.1016/J.TINS.2013.07.001.
74. Milton, Katharine. 1981. “Distribution Patterns of Tropical Plant Foods as an Evolutionary Stimulus to Primate Mental Development.” *American Anthropologist* 83 (3): 534–48. https://doi.org/10.1525/AA.1981.83.3.02A00020.
75. Mink, J. W., R. J. Blumenschine, and D. B. Adams. 1981. “Ratio of Central Nervous System to Body Metabolism in Vertebrates: Its Constancy and Functional Basis.” *Https://Doi.Org/10.1152/Ajpregu.1981.241.3.R203* 10 (2): 203–12. https://doi.org/10.1152/AJPREGU.1981.241.3.R203.
76. Morand-Ferron, Julie, Daniel Sol, and Louis Lefebvre. 2007. “Food Stealing in Birds: Brain or Brawn?” *Animal Behaviour* 74 (6): 1725–34. https://doi.org/10.1016/J.ANBEHAV.2007.04.031.
77. Navarrete, Ana, Carel P. van Schaik, and Karin Isler. 2011. “Energetics and the Evolution of Human Brain Size.” *Nature* 480 (7375): 91–93. https://doi.org/10.1038/nature10629.
78. Neves, Ben Hur Souto, Gabriel Palermo Del Rosso Barbosa, Ana Carolina de Souza Rosa, Steffanie Severo Picua, Gabriela Mendes Gomes, Priscila Marques Sosa, and Pâmela Billig Mello-Carpes. 2020. “On the Role of the Dopaminergic System in the Memory Deficits Induced by Maternal Deprivation.” *Neurobiology of Learning and Memory* 173 (September): 107272. https://doi.org/10.1016/J.NLM.2020.107272.
79. Niven, Jeremy E., and Simon B. Laughlin. 2008. “Energy Limitation as a Selective Pressure on the Evolution of Sensory Systems.” *The Journal of Experimental Biology* 211 (Pt 11): 1792–1804. https://doi.org/10.1242/JEB.017574.
80. Olkowicz, Seweryn, Martin Kocourek, Radek K. Luèan, Michal Porteš, W. Tecumseh Fitch, Suzana Herculano-Houzel, and Pavel Nemec. 2016. “Birds Have Primate-like Numbers of Neurons in the Forebrain.” *Proceedings of the National Academy of Sciences of the United States of America* 113 (26): 7255–60. https://doi.org/10.1073/PNAS.1517131113/SUPPL_FILE/PNAS.1517131113.SD03.DOCX.
81. Overington, Sarah E., Julie Morand-Ferron, Neeltje J. Boogert, and Louis Lefebvre.

2009. “Technical Innovations Drive the Relationship between Innovativeness and Residual Brain Size in Birds.” *Animal Behaviour* 78 (4): 1001–10. https://doi.org/10.1016/J.ANBEHAV.2009.06.033.

82. Parker, Sue Taylor, and Kathleen R. Gibson. 1977. “Object Manipulation, Tool Use and Sensorimotor Intelligence as Feeding Adaptations in Cebus Monkeys and Great Apes.” *Journal of Human Evolution* 6 (7): 623–41. https://doi.org/10.1016/S0047-2484(77)80135-8.
83. Pérez-Barbería, F. Javier, Susanne Shultz, and Robin I.M. Dunbar. 2007. “Evidence for Coevolution of Sociality and Relative Brain Size in Three Orders of Mammals.” *Evolution* 61 (12): 2811–21. https://doi.org/10.1111/j.1558-5646.2007.00229.x.
84. Pika, Simone, Miriam Jennifer Sima, Christian R. Blum, Esther Herrmann, and Roger Mundry. 2020. “Ravens Parallel Great Apes in Physical and Social Cognitive Skills.” *Scientific Reports* 10 (1). https://doi.org/10.1038/S41598-020-77060-8.
85. Pol, Anthony N. Van Den, Jean Pierre Wuarin, and F. Edward Dudek. 1990. “Glutamate, the Dominant Excitatory Transmitter in Neuroendocrine Regulation.” *Science* 250 (4985): 1276–78. https://doi.org/10.1126/SCIENCE.1978759.
86. Polymeropoulos, Elias T., Rebecca Oelkrug, and Martin Jastroch. 2018. “Editorial: The Evolution of Endothermy–From Patterns to Mechanisms.” *Frontiers in Physiology* 9 (July). https://doi.org/10.3389/fphys.2018.00891.
87. Pontzer, Herman, Mary H Brown, David A Raichlen, Holly Dunsworth, Brian Hare, Kara Walker, Amy Luke, et al. 2016. “Metabolic Acceleration and the Evolution of Human Brain Size and Life History.” *Nature* 533 (7603): 390–92. https://doi.org/10.1038/nature17654.
88. Puścian, Alicja, Szymon Lęski, Tomasz Górkiewicz, Ksenia Meyza, Hans-Peter Lipp, and Ewelina Knapska. 2014. “A Novel Automated Behavioral Test Battery Assessing Cognitive Rigidity in Two Genetic Mouse Models of Autism.” *Frontiers in Behavioral Neuroscience* 8: 140. https://doi.org/10.3389/fnbeh.2014.00140.
89. Raichlen, David A., and Adam D. Gordon. 2011. “ReRaichlen, D. A., & Gordon, A. D. (2011). Relationship between Exercise Capacity and Brain Size in Mammals. PLoS ONE, 6(6), E20601. Https://Doi.Org/10.1371/Journal.Pone.0020601lationship between Exercise Capacity and Brain Size in Mammals.” Edited by Andrew Iwaniuk. *PLoS ONE* 6 (6): e20601. https://doi.org/10.1371/journal.pone.0020601.
90. Rebola, Nelson, Mario Carta, and Christophe Mulle. 2017. “Operation and Plasticity of Hippocampal CA3 Circuits: Implications for Memory Encoding.” *Nature Reviews*

Neuroscience 2017 18:4 18 (4): 208–20. https://doi.org/10.1038/nrn.2017.10.

91. Rhodes, Justin S, Susan Jeffrey, Isabelle Girard, Gordon S Mitchell, Theodore Garland, and Fred H Gage. 2003. "Exercise Increases Hippocampal Neurogenesis to High Levels but Does Not Improve Spatial Learning in Mice Bred for Increased Voluntary Wheel Running." https://doi.org/10.1037/0735-7044.117.5.1006.
92. Roth, Gerhard, and Ursula Dicke. 2005. "Evolution of the Brain and Intelligence." *Trends in Cognitive Sciences* 9 (5): 250–57. https://doi.org/10.1016/J.TICS.2005.03.005.
93. Rudy, Jerry W., and Randall C. O'Reilly. 2001. "Conjunctive Representations, the Hippocampus, and Contextual Fear Conditioning." *Cognitive, Affective, & Behavioral Neuroscience 2001 1:1* 1 (1): 66–82. https://doi.org/10.3758/CABN.1.1.66.
94. Runge, Karen, Carlos Cardoso, and Antoine de Chevigny. 2020. "Dendritic Spine Plasticity: Function and Mechanisms." *Frontiers in Synaptic Neuroscience* 12 (August). https://doi.org/10.3389/FNSYN.2020.00036.
95. Rustay, Nathan, Kaitlin Browman, and Peter Curzon. 2008. "Cued and Contextual Fear Conditioning for Rodents." In , 19–37. https://doi.org/10.1201/noe1420052343.ch2.
96. Ruszczycki, Błażej, Zsuzsanna Szepesi, Grzegorz M Wilczynski, Monika Bijata, Katarzyna Kalita, Leszek Kaczmarek, and Jakub Wlodarczyk. 2012. "Sampling Issues in Quantitative Analysis of Dendritic Spines Morphology." *BMC Bioinformatics* 13: 213. https://doi.org/10.1186/1471-2105-13-213.
97. Sadowska, Joanna, Andrzej Gębczyński, and Marek Konarzewski. 2017. "Selection for High Aerobic Capacity Has No Protective Effect against Obesity in Laboratory Mice." *Physiology and Behaviour* 175: 130–36. https://doi.org/10.1016/j.physbeh.2017.03.034.
98. Sadowska, Julita, Andrzej K. Gębczyński, and Marek Konarzewski. 2013. "Basal Metabolic Rate Is Positively Correlated with Parental Investment in Laboratory Mice." *Proceedings of the Royal Society B: Biological Sciences* 280 (1753): 20122576. https://doi.org/10.1098/rspb.2012.2576.
99. Schaik, Carel P. Van, Zegni Triki, Redouan Bshary, and Sandra A. Heldstab. 2021. "A Farewell to the Encephalization Quotient: A New Brain Size Measure for Comparative Primate Cognition." *Brain, Behavior and Evolution* 96 (1): 1–12. https://doi.org/10.1159/000517013.
100. Schinazi, Victor R., Daniele Nardi, Nora S. Newcombe, Thomas F. Shipley, and

Russell A. Epstein. 2013. “Hippocampal Size Predicts Rapid Learning of a Cognitive Map in Humans.” *Hippocampus* 23 (6): 515–28. https://doi.org/10.1002/HIPO.22111.

101. Shulman, Robert G., Douglas L. Rothman, Kevin L. Behar, and Fahmeed Hyder. 2004. “Energetic Basis of Brain Activity: Implications for Neuroimaging.” *Trends in Neurosciences* 27 (8): 489–95. https://doi.org/10.1016/J.TINS.2004.06.005.
102. Simmen, Bruno, Patrick Pasquet, Shelly Masi, Georgius J. A. Koppert, Jonathan C. K. Wells, and Claude Marcel Hladik. 2017. “Primate Energy Input and the Evolutionary Transition to Energy-Dense Diets in Humans.” *Proceedings of the Royal Society B: Biological Sciences* 284 (1856): 20170577. https://doi.org/10.1098/rspb.2017.0577.
103. Smaers, J. B., R. S. Rothman, D. R. Hudson, A. M. Balanoff, B. Beatty, D. K.N. Dechmann, D. de Vries, et al. 2021. “The Evolution of Mammalian Brain Size.” *Science Advances* 7 (18): 9. https://doi.org/10.1126/SCIADV.ABE2101/SUPPL_FILE/ABE2101_SM.PDF.
104. Sobrero, Raúl, Laura J. May-Collado, Ingi Agnarsson, and Cristián E. Hernández. 2011. “Expensive Brains: ‘Brainy’ Rodents Have Higher Metabolic Rate.” *Frontiers in Evolutionary Neuroscience* 3 (JUL): 2. https://doi.org/10.3389/FNEVO.2011.00002/XML/NLM.
105. Sol, Daniel. 2008. “Revisiting the Cognitive Buffer Hypothesis for the Evolution of Large Brains.” *Biology Letters* 5 (1): 130–33. https://doi.org/10.1098/RSBL.2008.0621.
106. Sol, Daniel, Seweryn Olkowicz, Ferran Sayol, Martin Kocourek, Yicheng Zhang, Lucie Marhounová, Christin Osadnik, et al. 2022. “Neuron Numbers Link Innovativeness with Both Absolute and Relative Brain Size in Birds.” *Nature Ecology & Evolution 2022*, July, 1–9. https://doi.org/10.1038/s41559-022-01815-x.
107. Sonnay, Sarah, João M.N. Duarte, Nathalie Just, and Rolf Gruetter. 2015. “Compartmentalised Energy Metabolism Supporting Glutamatergic Neurotransmission in Response to Increased Activity in the Rat Cerebral Cortex: A 13C MRS Study in Vivo at 14.1 T.” *Journal of Cerebral Blood Flow and Metabolism* 36 (5): 928–40. https://doi.org/10.1177/0271678X16629482.
108. Sonnay, Sarah, Jordan Poirot, Nathalie Just, Anne Catherine Clerc, Rolf Gruetter, Gregor Rainer, and João M.N. Duarte. 2018. “Astrocytic and Neuronal Oxidative Metabolism Are Coupled to the Rate of Glutamate–Glutamine Cycle in the Tree Shrew Visual Cortex.” *Glia* 66 (3): 477–91. https://doi.org/10.1002/GLIA.23259.
109. Street, Sally E., Ana F. Navarrete, Simon M. Reader, and Kevin N. Laland. 2017.

"Coevolution of Cultural Intelligence, Extended Life History, Sociality, and Brain Size in Primates." *Proceedings of the National Academy of Sciences* 114 (30): 7908–14. https://doi.org/10.1073/pnas.1620734114.

110. Swallow, John G., Jack P. Hayes, Pawel Koteja, and Theodore Garland. 2009. "Selection Experiments and Experimental Evolution of Performance and Physiology." *Experimental Evolution: Concepts, Methods, and Applications of Selection Experiments*, December, 301–51. https://doi.org/10.1525/CALIFORNIA/9780520247666.003.0012.
111. Thurber, Caitlin, Lara R. Dugas, Cara Ocobock, Bryce Carlson, John R. Speakman, and Herman Pontzer. 2019. "Extreme Events Reveal an Alimentary Limit on Sustained Maximal Human Energy Expenditure." *Science Advances* 5 (6). https://doi.org/10.1126/SCIADV.AAW0341/SUPPL_FILE/AAW0341_SM.PDF.
112. Trachtenberg, Joshua T., Brian E. Chen, Graham W. Knott, Guoping Feng, Joshua R. Sanes, Egbert Welker, and Karel Svoboda. 2002. "Long-Term in Vivo Imaging of Experience-Dependent Synaptic Plasticity in Adult Cortex." *Nature 2003 420:6917* 420 (6917): 788–94. https://doi.org/10.1038/nature01273.
113. Tsuboi, Masahito, Wouter van der Bijl, Bjørn Tore Kopperud, Johannes Erritzøe, Kjetil L. Voje, Alexander Kotrschal, Kara E. Yopak, Shaun P. Collin, Andrew N. Iwaniuk, and Niclas Kolm. 2018. "Breakdown of Brain–Body Allometry and the Encephalization of Birds and Mammals." *Nature Ecology & Evolution* 2 (9): 1492–1500. https://doi.org/10.1038/s41559-018-0632-1.
114. Voikar, Vootele, Giovanni Colacicco, Oliver Gruber, Elisabetta Vannoni, Hans-Peter Lipp, and David P. Wolfer. 2010. "Conditioned Response Suppression in the IntelliCage: Assessment of Mouse Strain Differences and Effects of Hippocampal and Striatal Lesions on Acquisition and Retention of Memory." *Behavioural Brain Research* 213 (2): 304–12. https://doi.org/10.1016/j.bbr.2010.05.019.
115. Whiten, Andrew, and Carel P. Van Schaik. 2007. "The Evolution of Animal Cultures and Social Intelligence." *Philosophical Transactions of the Royal Society B: Biological Sciences* 362 (1480): 603–20. https://doi.org/10.1098/RSTB.2006.1998.
116. Withers, P. C. 1977. "Measurement of VO2, VCO2, and Evaporative Water Loss with a Flow-through Mask." *Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology* 42 (1): 120–23. https://doi.org/10.1152/JAPPL.1977.42.1.120.
117. Woerden, Janneke T. Van, Carel P. Van Schaik, and Karin Isler. 2010. "Effects of

Seasonality on Brain Size Evolution: Evidence from Strepsirrhine Primates." *American Naturalist* 176 (6): 758–67. https://doi.org/10.1086/657045/0.

118. Woerden, Janneke T. van, Erik P. Willems, Carel P. van Schaik, and Karin Isler. 2012. "Large Brains Buffer Energetic Effects of Seasonal Habitats in Catarrhine Primates." *Evolution; International Journal of Organic Evolution* 66 (1): 191–99. https://doi.org/10.1111/J.1558-5646.2011.01434.X.
119. Wong-Riley, Margaret. 1979. "Changes in the Visual System of Monocularly Sutured or Enucleated Cats Demonstrable with Cytochrome Oxidase Histochemistry." *Brain Research* 171 (1): 11–28. https://doi.org/10.1016/0006-8993(79)90728-5.
120. Wong-Riley, Margaret T.T. 2012. "Bigenomic Regulation of Cytochrome c Oxidase in Neurons and the Tight Coupling between Neuronal Activity and Energy Metabolism." *Advances in Experimental Medicine and Biology* 748: 283–304. https://doi.org/10.1007/978-1-4614-3573-0_12/FIGURES/3.
121. Wrangham, Richard, and Nancy Lou Conklin-Brittain. 2003. "Cooking as a Biological Trait." *Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology* 136 (1): 35–46. https://doi.org/10.1016/S1095-6433(03)00020-5.
122. Yuste, R., and T. Bonhoeffer. 2001. "Morphological Changes in Dendritic Spines Associated with Long-Term Synaptic Plasticity." *Annual Review of Neuroscience* 24: 1071–89. https://doi.org/10.1146/ANNUREV.NEURO.24.1.1071.
123. Zafon, C. 2007. "Oscillations in Total Body Fat Content through Life: An Evolutionary Perspective." *Obesity Reviews* 8 (6): 525–30. https://doi.org/10.1111/J.1467-789X.2007.00377.X.
124. Zhang, Xiaomin, Alois Schlögl, and Peter Jonas. 2020. "Selective Routing of Spatial Information Flow from Input to Output in Hippocampal Granule Cells." *Neuron* 107 (6): 1212-1225.e7. https://doi.org/10.1016/J.NEURON.2020.07.006.

10 Supplementary materials

SAS codes

Exemplary SAS code used for analyses of BMR. Similarly structured codes were used for the analyses of anatomic traits.

```
BMR=log(BMR);
BM=log(BM);
title '';
run;
proc sort;
by line subline;
run;
proc print;
var line subline Setup channel BM BMR;
run;
  proc univariate data=BMR normaltest robustscale plot;
var  BMR;
by line;
run;
proc mixed covtest;
class line  subline Setup channel;
model BMR=  setup channel line BM line*BM /solution ;
random subline(line);
lsmeans line/tdiff;
run;
proc mixed covtest;
class line  subline Setup channel;
model BMR=  Setup channel line BM /solution ;
random subline(line);
lsmeans line/tdiff;
run;
proc mixed covtest;
class line  subline ;
model BMR=  line BM /solution ;
random subline(line);
lsmeans line/tdiff;
run;
proc mixed covtest;
class line  subline ;
model BMR=  line BM /solution ;
*random subline(line);
lsmeans line/tdiff;
run;
quit;
```

Exemplary SAS code used for the analysis of the number of nosepokes (i.e. correct responses) in the reward-seeking discrimination learning task. Similarly structured code was used for the analyses of other behavioural tasks.

```
INPUT ID$ line$ subline dark_C1 light_C1 dark_C2 light_C2 dark_rog1
light_rog1 dark_rog2 light_rog2 Generation batch;
* The above variables are defined in the data file deposited in DRYAD www;
  DATA long ;
  SET Wide;
*Below the data are re-structured to conform to the Proc Mixed format;
  correct = (dark_C1); rog = (dark_rog1-dark_C1); phase = 'dark'; time = 1;
b=batch;f=generation; period=time; OUTPUT;
  correct = (light_C1); rog = (light_rog1-light_C1); phase = 'light'; time
= 1; b=batch; f=generation; period=time; OUTPUT;
  correct = (dark_C2); rog = (dark_rog2-dark_C2); phase = 'dark'; time = 2;
b=batch; f=generation; period=time; OUTPUT;
  correct = (light_C2); rog = (light_rog2-light_C2); phase = 'light'; time
= 2; b=batch; f=generation; period=time; OUTPUT;
  DROP dark_C1 light_C1 dark_C2 light_C2 dark_rog1 light_rog1 dark_rog2
light_rog2 generation batch;
RUN;
 PROC PRINT DATA=long ;
RUN;
  PROC MIXED DATA=long covtest;
  CLASS ID line subline phase time;
   MODEL correct =  line  time phase time*phase line*phase line*time rog /
solution;
   random subline(line);
   repeated time*phase / subject=id type=VC;
   contrast "HBMR vs rest" line*time  3 -3 -1 1 -1 1 -1 1/;
   contrast "LBMR vs rest" line*time  -1 1 3 -3 -1 1 -1 1/;
   contrast "RB vs rest" line*time  -1 1 -1 1 -1 1 3 -3/;
     LSMEANS line/ tdiff;
   LSMEANS line*time/ tdiff;
 run;
 PROC MIXED DATA=long;
  CLASS ID line subline phase time;
   MODEL correct =  line  time phase time*phase line*phase line*time rog /
solution;
   *random subline(line)/;
   repeated time*phase / subject=id type=VC;
   LSMEANS line/ tdiff;
   LSMEANS line*time/ tdiff;
 run;
  *ods pdf close;
quit;
```

Raw images acquired with microscopy were analyzed with ImageJ software (NIH):

- GFAP (astrocyte) immunoreactivity in hippocampus (ROI) was assessed according to listed ImageJ macro:

```
run("Subtract Background...", "rolling=10");
run("Auto Local Threshold", "method=Phansalkar radius=25 parameter_1=0
parameter_2=0");
run("Options...", "iterations=2 count=6 do=Close");
run("Analyze Particles...", "size=50-Infinity summarize");
```