The evolution of nesting behaviour in *Peromyscus* mice

Caitlin L. Lewarch and Hopi E. Hoekstra

Department of Organismic and Evolutionary Biology,

Department of Molecular and Cellular Biology,

Museum of Comparative Zoology, Howard Hughes Medical Institute, Harvard University, Cambridge,

MA 02138 USA

Corresponding Author:
Hopi Hoekstra
Biological Labs
16 Divinity Ave
Cambridge, MA 02138
(617) 496-9040
hoekstra@oeb.harvard.edu

ABSTRACT

Structures built by animals, such as nests, often can be considered extended phenotypes that facilitate the study of animal behaviour. For rodents, nest building is both an important form of behavioural thermoregulation and a critical component of parental care. Changes in nest structure or the prioritization of nesting behaviour are therefore likely to have consequences for survival and reproduction, and both biotic and abiotic environmental factors are likely to influence the adaptive value of such differences. Here we first develop a novel assay to investigate interspecific variation in the nesting behaviour of deer mice (genus *Peromyscus*). Using this assay, we find that, while there is some variation in the complexity of the nests built by *Peromyscus* mice, differences in the latency to begin nest construction are more striking. Four of the seven taxa examined here build nests within an hour of being given nesting material, but this latency to nest is not related to ultimate differences in nest structure, suggesting that the ability to nest is relatively conserved within the genus, but species differ in their prioritization of nesting behaviour. We also find that latency to nest is not correlated with body size, climate, or the construction of burrows that create microclimates. However, the four taxa with short nesting latencies all have monogamous mating systems, suggesting that differences in nesting latency may be related to social environment. This detailed characterization of nesting behaviour within the genus provides an important foundation for future studies of the genetic and neurobiological mechanisms that contribute to the evolution of behaviour.

Key words: behavioural evolution, comparative method, deer mice, extended phenotype

Running head: Evolution of nesting behaviour

INTRODUCTION

Animal architectures – from the webs spun by spiders to the dams built by beavers – can both facilitate the study of behaviour and provide insight into the selective forces at that act on behavioural variation (Hansell, 1984, 2005). Such structures can be considered "extended phenotypes," or traits influenced by genetics but *extended* outside the body of the individual organism (Dawkins, 1982). Building behaviours are often innate and species-specific; for example, the resulting structures have been used for classification purposes in insects and some birds (Hansell, 1984; Knerer et al., 2012; Schmidt, 1964; Winkler et al., 1993). These structures reflect stereotyped patterns of behaviour and the neural circuits that generate these motor patterns, allowing us to study behaviour and the nervous system by proxy. Moreover, the structures themselves serve important functions and can confer readily quantifiable fitness benefits on the animals that construct them (Hayward, 1965; Mainwaring et al., 2014; Sealander, 1952).

A widespread and important type of building behaviour is the collection and processing of environmental materials to produce a nest. Nests serve a wide variety of purposes for the animals that construct them. For small-bodied animals, such as rodents, nests provide insulation and reduce the energy expended on the maintenance of body temperature (Pearson, 1960; Sealander, 1952; Vogt & Lynch, 1982). In animals with altricial young, like many birds and rodents, nests are especially critical to protect offspring from heat loss and predation (Bult et al., 1997; Collias, 1964; Lynch & Possidente, 1978; Southwick, 1955). The nest may even serve as a catalyst for social behaviour — nest and bower construction can be integral to courtship in birds (Mainwaring et al., 2014), and investment in elaborate nests likely has been instrumental in the evolution of eusociality in insects (Hansell, 2005). Depending on the species in question and the environment in which they live, nests may be built in trees, in pre-existing cavities, or in burrow systems that are also constructed by the animal (Collias, 1964; Dooley & Dueser, 1990; Weber & Hoekstra, 2009). While the excavation of burrows is itself a type of animal architecture, nests are often separate structures, made by collecting and processing vegetation and other material from an animal's environment.

Both the structure of a completed nest and the timing of nest building may be relevant traits for natural selection, and each has distinct implications for the proximate and ultimate factors that contribute to behavioural differences among taxa. Variation in nest structure, as is observed in birds, suggests that animals may differ either in their ability to construct nests or in the desired properties of their nests (Mainwaring et al., 2014). At the level of proximate mechanism, variation could result from morphological differences in the animals, fundamental changes in their stereotyped motor patterns, or changes in a more abstract encoding of the animal's target structure. Moreover, variation in nest structures suggests that the characteristics of the nest itself have fitness consequences. Prime examples of such relationships include the pendulous entrances of some weaverbird nests, which are protective against snake predation (Collias, 1964; Crook, 1963), or the increased size and weight of robin, warbler, and finch nests built at colder northern versus southern latitudes (Crossman et al., 2011). Variation in the timing of nesting behaviour, on the other hand, implies that animals differ in their motivation to engage in otherwise conserved behavioural patterns, and suggests that the prioritization of nesting relative to other elements of the animal's behavioural repertoire is relevant for selection. Prioritization can occur at different scales, from time invested over the course of a single night to relative time spent on the behaviour during different seasons. As the collection of nesting material can be energetically costly and expose the animal to predation (Collias, 1964; Mainwaring et al., 2014), it may be beneficial for an animal to prioritize other behaviours in environmental conditions where heat loss, for example, is not a pressing concern. While population differences in nest size have been studied within and between species of rodents (King et al., 1964; Lynch, 1992), we do not know how the prioritization of nesting behaviour has evolved.

To determine how and why these features of nesting behaviour evolve, we focused on deer mice (genus *Peromyscus*), which have adapted to a wide range of habitats and microhabitats across North America (Bedford & Hoekstra, 2015; Blair, 1950; Dewey & Dawson, 2001). Specifically, deer mice live in climates with pronounced differences in winter temperatures (King et al., 1964), vary in body size, a trait associated with adaptation to cold in other rodents (Lynch, 1992), and have distinct social behaviour and parental care (Jašarević et al., 2013; Turner et al., 2010), all of which may affect nest-building behaviour. Importantly, while these species have evolved in different environments, laboratory colonies allow us to perform behavioural experiments under carefully controlled conditions using animals that share a common environment (Bedford & Hoekstra, 2015). This is therefore an opportunity to explore the evolutionary consequences of different environmental parameters on heritable variation in nest-building behaviour.

Here we develop a novel behavioural assay to evaluate natural variation in both ability and motivation to nest in seven species and subspecies of *Peromyscus* mice. This detailed characterization of thermoregulatory nesting behaviour then provides a foundation to understand the evolution of this behaviour in natural populations.

METHODS

Ethical Note

All experimental procedures were approved by the Harvard University Institutional Animal Care and Use Committee. The animal housing facility in which these tests were performed maintains full AAALAC accreditation.

Experimental Cohort

We selected adult, reproductively inexperienced animals of both sexes from seven laboratory colonies of *Peromyscus,* representing five species, with well-characterized ecology and social systems (Table 1). While these colonies were isolated from natural populations (brought in from the wild between 2 and 71 years ago, depending on strain; Table 1), all animals in this study were born in captivity.

Animal Husbandry

All animals were bred and maintained under the same controlled conditions. We kept the animal housing rooms on a 16:8 LD cycle at 22°C. We housed animals in ventilated polysulfone mouse cages (Allentown, NJ) of standard size (19.7cm wide x 30.5cm long x 16.5cm high), with the exception of the *P. californicus* animals, which were housed in rat cages (28.6cm wide x 39.4cm long x 19.3cm high) due to their large body size (Allentown, NJ). For ordinary housing, we provided all cages with 2.5g of compressed cotton “Nestlet” (Ancare, Bellmore, NY), 8-10g folded paper “Enviro-Dri” nesting material (Shepherd Specialty Papers, Watertown, TN), a 0.6cm layer of Anderson’s Bed-o-cob (The Andersons, Inc., Maumee, OH), and enrichment consisting of a red polycarbonate (9.5cm x 4.8cm x 7.6cm) mouse hut (BioServ, Flemington, NJ) or a 15.2cm x 7.6cm inside diameter rat tunnel for the large *P. californicus* animals (BioServ, Flemington, NJ). All animals had *ad libitum* access to water and irradiated LabDiet Prolab Isopro RMH 3000 5P75 (LabDiet, St. Louis, MO). We socially housed animals in groups of 2-5 by species and sex after weaning (23 days for most species, 30 days for *P. californicus*), then tested them as adult virgins, averaging 2-6 months old (Table 1).

Behavioural Paradigm

<u>Standard Behavioural Assay</u>: Nesting behaviour in rodents is often assessed by measuring the weight of nesting material an animal uses over 24 hours (Hartung & Dewsbury, 1979; King et al., 1964;

Table 1: Experimental Cohort

Species (common name)	County Isolated	Year in Captivity*	Sample Size **total** (males, females)	Avg. Weight, Males (g ± sd)	Avg. Weight, Females (g ± sd)	Avg. Age (days ± sd)
P. maniculatus nubiterrae (cloudland deer mouse)	Westmoreland County, PA	2010	**47** (31,16)	18.7 ± 2.3	15.8 ± 2.8	164 ± 184
P. maniculatus bairdii (deer mouse)	Washtenaw County, MI	1946-1948	**95** (62,33)	20.3 ± 3.5	16.9 ± 1.6	106 ± 50
P. polionotus subgriseus (oldfield mouse)	Marion County, FL	1952	**130** (80,50)	14.3 ± 1.9	15.5 ± 1.7	107 ± 57
P. polionotus leucocephalus (Santa Rosa Island beach mouse)	Okaloosa County, FL	2015	**37** (23,14)	14.2 ± 1.1	14.4 ± 2.7	71 ± 17
P. leucopus (white-footed mouse)	Avery County, NC	1982-1985	**35** (22,13)	21.7 ± 4.1	20.1 ± 2.7	66 ± 7
P. gossypinus (cotton mouse)	Jackson County, FL	2009	**27** (19,8)	25.2 ± 7.0	21.9 ± 3.3	72 ± 9
P. californicus (California mouse)	Ventura County, CA	1979-1987	**48** (25,23)	42.1 ± 5.2	41.5 ± 7.0	126 ± 30

*Some species were brought into captivity multiple times over several years, see (Bedford & Hoekstra., 2015). Female *maniculatus bairdii* animals give birth to their first litter when they are approximately three months old (Bedford & Hoekstra, 2015), and generation times for other species are similar in the lab.

Layne, 1969; Lynch & Hegmann, 1973), which is readily quantifiable but can obscure variation in the timing of the behaviour or the structure of the nests the animals construct. To measure these aspects of nesting behaviour, we designed a novel assay that consists of an overnight habituation period followed by three consecutive days of testing. On the day before a trial began, we weighed and singly housed adult virgin animals in new mouse cages (including *P. californicus*) with 5g of compressed cotton nesting material (or two "nestlets", see above), 0.6cm layer of Anderson's Bed-o-cob, and a red polycarbonate mouse hut. On the morning following habituation to the novel cage, we took photos of the nest from up to three angles (top and two side views), then removed the mouse hut and replaced all cotton nesting material with 5g of fresh compressed cotton nestlet. The replacement of nesting material during these trials always occurred between 4.5 and 6.5 hours after the lights came on. At one hour after the replacement of nesting material, we again took photographs of the nest from multiple angles and added the mouse hut back to the cage. We repeated this process on the following two mornings for a total of three sets of photographs (day 1, day 2, and day 3) at each of the two time points (1h and overnight). Research assistants blinded to the species and sex of the animal later scored these nest photographs according to a standardized scale (Fig. 1; Supplemental Table S1). Scores ranged from 0 (no visible shredding) to 4 (a full "dome" nest with overhead coverage) with only full and half scores given.

<u>Increased Nesting Material</u>: To examine whether the amount of nesting material had an impact on nest scores in the largest species (*P. californicus*; approximately 42g, on average), we modified the nesting experiment in two ways. First, we singly housed an independent cohort of 21 adult *P. californicus* animals as above, but provided them with an increasing amount of cotton nesting material on four consecutive days: 5g on day 1, 10g on day 2, 15g on day 3, and 20g on day 4. We photographed nests and exchanged cotton nesting material once every 24 hours, and a research assistant blind to experimental conditions scored these photographs as above to establish whether this increase was

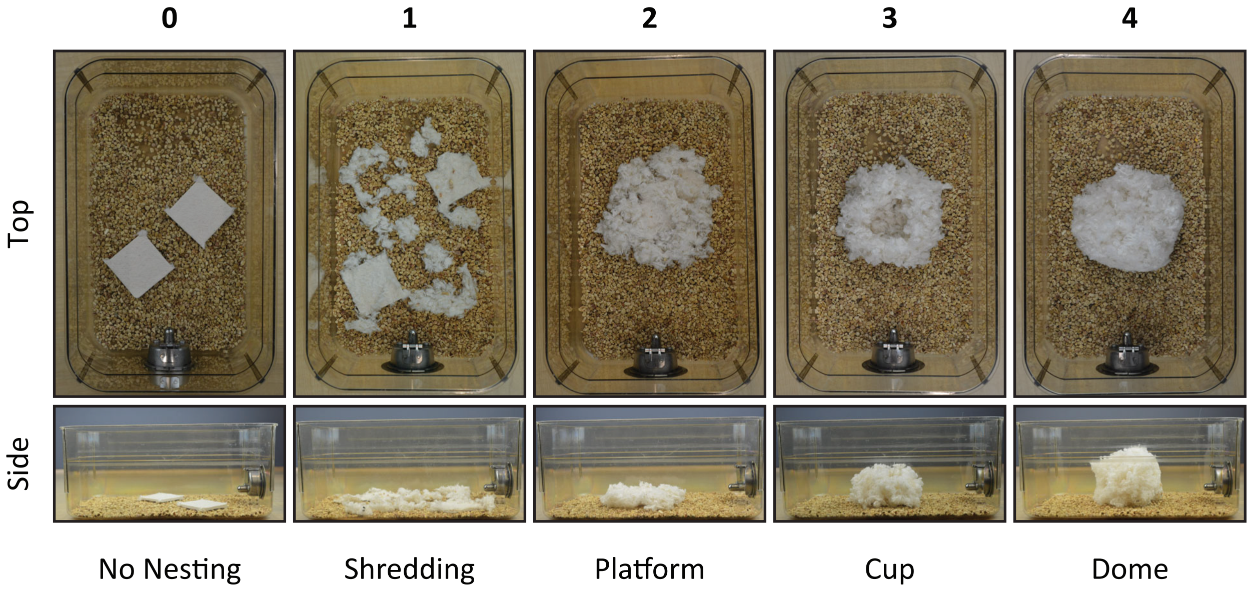


Figure 1: Nest scoring scale. Nests are scored on a scale from 0 (no manipulation of the nesting material) to 4 (a full cotton "dome") in increments of 0.5. Representative nests for each of the five integer scores are shown from both the top and side view, and a brief descriptor is provided for each. Classification is according to criteria provided in Supplementary Table S1.

sufficient to alter overnight nest scores. Based on the results of these experiments, we then assayed an independent group of 23 adult *P. californicus* animals to evaluate their overnight nesting behaviour using 20g of cotton nesting material in an otherwise standard nesting assay (see "Standard Behavioural Assay" above).

Climate Data

We drew average winter (December/January/February) temperature data from National Oceanic and Atmospheric Administration (NOAA) 30-year climate normals (Arguez et al., 2010), and averaged these data by state or county of origin for each colony (Table 2).

Data Analysis

We performed statistical analyses in R using non-parametric methods for the ordinal nest scores. We summarized an animal's behaviour across the three trial days by its median score (to reflect central tendency) or its maximum score (to represent best effort) at each time point after the replacement of nesting material. To identify differences between groups, we used Kruskal-Wallis rank-sum tests, and then used Bonferroni-corrected Wilcoxon rank-sum tests for subsequent pairwise comparisons between species or sexes. For the experiment that tests the effects of increased access to nesting material in a cohort of *P. californicus* animals, we used a Friedman rank-sum test.

For comparative analyses, we first generated an ultrametric tree using Grafen's method (Grafen, 1989; Symonds & Blomberg, 2014) and the known topology of the species relationships (Fig. 2A; Bedford & Hoekstra, 2015; Bradley et al., 2007; Weber & Hoekstra, 2009). To test for a relationship between species-level median 1h nest scores and species-level average weights, latitudes of origin, and winter temperatures at sites of origin, we performed phylogenetic generalized least squares (PGLS) analysis using the ape and nlme packages in R (Paradis et al., 2004; Pinheiro et al., 2017; Symonds & Blomberg,

Table 2: Environmental Context

Taxon	Latitude (°N)^A	Avg. Winter Temp. (°C) ^{B}	Habitat	Habitat Use	Habitat Ref.	Burrows^C	Nest Location	Mating System^D
P. m. nubiterrae	40.2	-1.2	forest	semi-arboreal	(Blair, 1950)	simple/short	arboreal; tree cavities (Wolff & Durr, 1986; Wolff & Hurlbutt, 1982)	M
P. m. bairdii	42.3	-3.4	prairie/grassland	terrestrial	(Blair, 1950)	simple/short	in burrows (Morris & Kendeigh, 1981)	P
P. p. subgriseus	29.2	14.4	sandy soil/grassland	semi-fossorial	(Blair, 1950)	complex/long	in burrows (Dawson et al., 1988)	M
P. p. leucocephalus	30.4	11.0	white sand beach	semi-fossorial	(Blair, 1950; Sumner, 1926)	complex/long	in burrows (Blair, 1951)	M
P. leucopus	36.1	0.9	deciduous forest	semi-arboreal	(Blair, 1950; Lackey et al., 1985)	simple/short	seasonally dependent; arboreal or ground nests (Wolff & Durr, 1986; Wolff & Hurlbutt., 1982)	P
P. gossypinus	30.7	11.3	hardwood forest, mesic hammocks, swamps	semi-arboreal	(Wolfe & Linzey, 1977)	no laboratory data	diverse but arboreal preferred; in or under logs and stumps, in tree cavities (Ivey, 1949; Klein & Layne, 1978; Wolfe & Linzey, 1977)	P
P. californicus	34.1	12.3	scrub/chaparral	semi-arboreal	(Clark, 1936; M'Closkey, 1976; Merritt, 1978; Meserve, 1977)	none	under logs, in woodrat (*Neotoma*) dens (Merritt, 1974, 1978)	M

A. Latitude of county of origin (Table 1).

B. 30-year NOAA climate normals for average Dec/Jan/Feb temperatures, pooled by county of origin for each colony (Arguez et al., 2010).

C. Burrows produced in a laboratory assay from (Weber & Hoekstra, 2009), with the exception of *P. m. nubiterrae* (Hu & Hoekstra, 2017).

D. Most likely mating system (Monogamous or Promiscuous) according to (Turner et al., 2010), with the exception of *P. m. nubiterrae* (Wolff & Cicirello, 1991) and *P. gossypinus* (Dewsbury et al., 1980; McCarley, 1959; Pearson, 1953).

2014). Covariance due to relatedness was modelled by Brownian motion using the corBrownian function in ape. The covariance was then included as a correlation parameter in the generalized least squares analyses in nlme. The effect of each environmental variable on 1h nest scores was tested independently. To test whether short nesting latency is dependent on other discrete traits (complex burrowing or mating system, as indicated in Table 2), we performed Pagel's binary character correlation test using the fitPagel function in the phytools package in R (Pagel, 1994; Revell, 2012). For this test, we utilized the fitMk method, allowed all rates of change to be different between states (model="ARD"), and set nesting latency (short vs. intermediate/long) to be dependent on the state of either mating system (monogamous vs. promiscuous) or burrow complexity (complex vs. simple/absent). As there are no laboratory data on burrowing behaviour in *P. gossypinus*, this species was excluded from the latter analysis.

RESULTS

Interspecific variation in nesting latency

To measure an animal's motivation to nest, we assayed individuals from seven *Peromyscus* taxa with known evolutionary relationships (Fig. 2A). First, we analysed the median of the three scores an animal received one hour after the replacement of nesting material, which reflects the tendency of the animal to begin nesting shortly after their nest is disturbed. Scores at 1h were significantly correlated across the three days in the full dataset (Spearman rank correlations: day 1 vs. day 2 $r_s = 0.75$, day 1 vs. day 3 $r_s = 0.68$, day 2 vs. day 3 $r_s = 0.78$, $N=419$, $P<2.2\text{x}10^{-16}$ for each), and species comparisons were largely the same whether three-day medians or maxima were used (see below). The median nest scores at 1h following the initiation of the trial varied dramatically among the taxa we assayed (Fig. 2B; Kruskal-Wallis test: $H_6= 216.85$, $P<2.2\text{x}10^{-16}$). Four taxa (*P. m. nubiterrae*, *P. p. subgriseus*, *P. p. leucocephalus*, and *P. californicus*) had high, statistically indistinguishable scores at the 1h time point (Supplemental

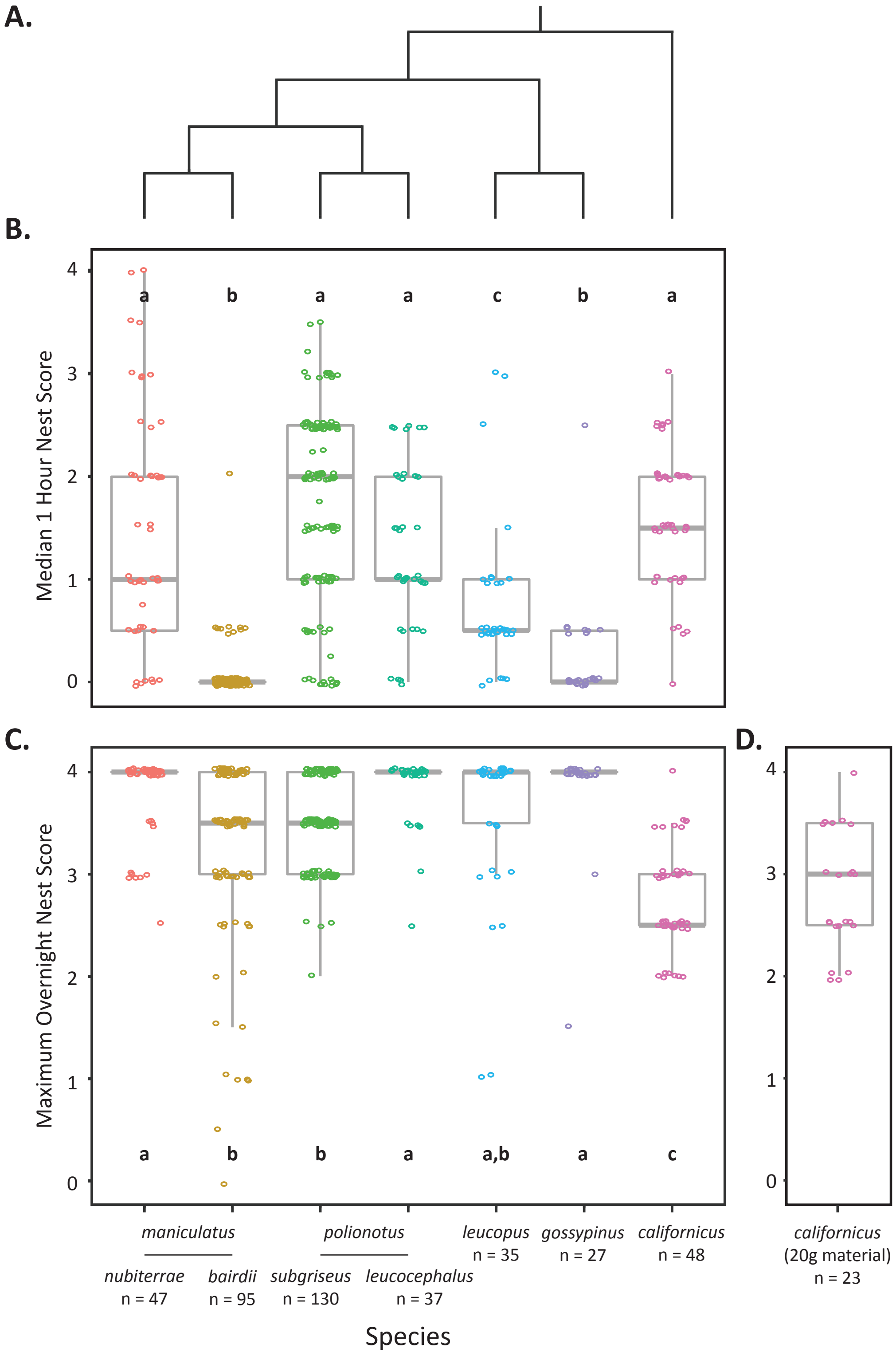


Figure 2: Interspecific differences in nesting behaviour. (A) Phylogenetic relationships among the *Peromyscus* taxa included in this study; modified from (Bedford & Hoekstra, 2015; Weber et al., 2009). (B) Median nest scores 1h after receiving new nesting material and (C) maximum overnight nest scores for each animal over the 3 trial days. Letters indicate species groups that do not significantly differ from one another, while all other pairwise comparisons are significant (Wilcoxon rank-sum test, Bonferroni-corrected *P*<0.05). (D) *P. californicus* animals given 20g of nesting material do not differ in maximum overnight scores from those given 5g of nesting material (Wilcoxon rank-sum test, *P*=0.47). Sample sizes are provided below.

Table S2, Wilcoxon rank-sum test, Bonferroni-corrected $P>0.05$ for each pairwise comparison), suggesting that they began to construct their nests relatively quickly and progressed past the point of just shredding the material. In addition, these four taxa differed significantly from the other three taxa we assayed (Supplemental Table S2, Wilcoxon rank-sum test, Bonferroni-corrected $p<0.05$ for each pairwise comparison). *P. leucopus* animals received intermediate scores that reflect a tendency to shred the material, but not arrange it into a nest, at the 1hr time point. These scores were significantly different from those received by all other species (Supplemental Table S2, Wilcoxon rank-sum test, Bonferroni-corrected $P<0.05$ for each pairwise comparison). Finally, *P. m. bairdii* and *P. gossypinus* animals had equivalently low scores (Supplemental Table S2, Wilcoxon rank-sum test, Bonferroni-corrected $P=0.29$), which indicate that they did not manipulate the nesting material in the first hour and suggest that they are relatively slow to initiate nest construction. Ranking of each taxon's performance largely followed the same pattern whether maximum or median nest scores were used (Supplemental Figure S1). The only exception was one species difference: while the median nest scores of *P. gossypinus* and *P. m. bairdii* animals at 1h were indistinguishable, *P. gossypinus* were slightly more likely to shred the nesting material on at least one of the trial days and therefore had slightly, but significantly, higher maximum scores (Wilcoxon rank-sum test: $W=839$, $N_1=27$, $N_2=95$, Bonferroni-corrected $P = 0.049$). We note that taxa differ in the variance of their nest scores: this likely results from within-species variation in nest-building efficiency, time spent nesting, and/or the precise initiation time during the first hour. In sum, based on our analysis of 1h median scores, we identified three main groups of nest builders in our assay: those with short, intermediate or long latencies to nest.

Interspecific differences in nesting ability

We next asked whether these taxa differed in their overall ability to construct a three-dimensional nest. To establish the highest-scoring nest that an animal was capable of producing, we used the maximum score achieved over the individual's three overnight time points, which represents the animal's best effort during the longest interval of the trial. Maximum overnight scores varied significantly among taxa (Fig. 2C; Kruskal-Wallis test: H_6=127.21, P<2.2x10^{-16}), although most animals built full or partial domes. The highest scoring nests were consistently constructed by *P. m. nubiterrae*, *P. p. leucocephalus*, *P. leucopus*, and *P. gossypinus* animals, which tended to build statistically indistinguishable full domes (Supplemental Table S2, Wilcoxon rank-sum test, Bonferroni-corrected *P*>0.05 for each pairwise comparison). Three taxa – *P. m. bairdii*, *P. p. subgriseus*, and *P. leucopus* – had equivalently high maximum scores (Supplemental Table S2, Wilcoxon rank-sum test, Bonferroni-corrected *P*>0.05 for each pairwise comparison), and *P. m. bairdii* and *P. p. subgriseus,* which tended to build domes with only partial cover, were significantly different from all but *P. leucopus* animals (Supplemental Table S2, Wilcoxon rank-sum test, Bonferroni-corrected *P*<0.05 for each pairwise comparison). Finally, *P. californicus* animals tended to build nests with walls but without overhead cover, and had significantly lower maximum nest scores than all other species tested (Supplemental Table S2, Wilcoxon rank-sum test, Bonferroni-corrected *P*<0.05 for each pairwise comparison). Notably, we found that all species had at least one individual who constructed a domed nest with full cover (maximum nest score, "4") during the assay, suggesting that all species are capable of building a "complete" nest if given enough time. However, some species showed a large variance in nest scores, and *P. californicus* tended to have lower maximum scores than the other species.

Nest-building behaviour in the large *P. californicus* mice

P. californicus animals are much larger than the other taxa included in this study (Table 1), and therefore might require more material to construct a dome nest with overhead cover. To test the possibility that these animals built lower-scoring nests because 5g of nestlet was an insufficient amount of nesting material, we conducted two additional experiments. First, we gave a group of *P. californicus* animals increasing amounts of nesting material on four consecutive days and evaluated the nests they produced in each 24-hour interval. We found that increasing nesting material from 5g to 20g could increase overnight nesting scores (Supplemental Fig. S2; Friedman test: X^2_3=13.468, p=0.004). However, when we provided an independent group of *P. californicus* animals with 20g of nesting material during a three-day trial (Fig. 2D), there was no difference in overnight maximum scores between those *P. californicus* given 5g of nestlet and those given 20g (Wilcoxon rank-sum test: W= 609, N_1=23, N_2=48, P=0.47). Moreover, the maximum overnight nest scores for *P. californicus* given 20g of nestlet remained significantly lower than the maximum nest scores for all other species (Supplemental Table S2, Wilcoxon rank-sum test, Bonferroni-corrected P<0.05 for each pairwise comparison). Thus, the poor nest construction of *P. californicus* in this assay cannot be attributed simply to insufficient nesting material relative to its large body size.

Sex differences in nesting

We next investigated whether there were any sex differences in the nest scores produced by each species and subspecies. Only two taxa showed evidence of sexual dimorphism in nesting (Fig. 3). Both male *P. m. nubiterrae* and male *P. p. subgriseus* animals built higher scoring nests than their female counterparts one hour after the start of the assay (Supplemental Table S3, Wilcoxon rank-sum test, Bonferroni-corrected P = 0.03 and 0.002, respectively), and *P. p. subgriseus* males also built higher scoring nests at the overnight time point (Supplemental Table S3, Supplemental Fig. S3; Wilcoxon rank-sum test, Bonferroni-corrected P = 0.008). No other species showed evidence of sex differences in nest scores at either time point (Supplemental Table S3, Wilcoxon rank-sum test, Bonferroni-corrected P >0.05 for each pairwise comparison). Therefore, while there was no sexual dimorphism in nesting behaviour for most taxa, in both instances when sex differences were observed, males constructed higher-scoring nests than the females.

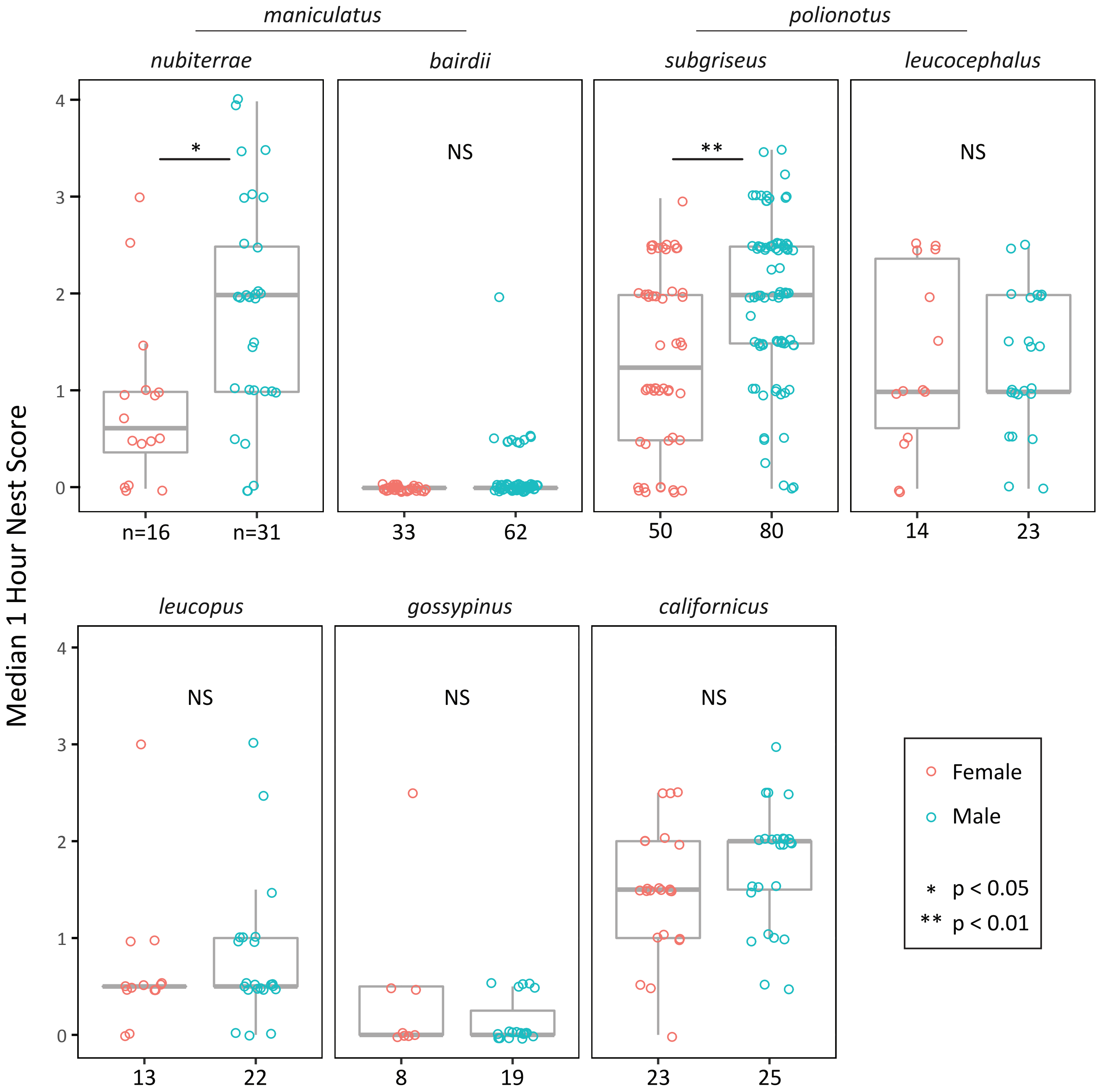


Figure 3: Sex differences in nesting latency. Sex-specific median 1h nest scores for each *Peromyscus* taxa tested. Significant sex differences in median nest score occurred in only two groups: *P. maniculatus nubiterrae* (Wilcoxon rank-sum test, Bonferroni-corrected *P*=0.03) and *P. polionotus subgriseus* (*P*=0.002). Sample sizes are provided below. NS = non-significant.

Association between body size and nest building

To determine whether body size had an effect on nest-building behaviour, we tested for correlations between weight and performance in the nesting assay. We found that weight significantly varied by species, sex, and species-by-sex interactions in our experimental cohort (two-way ANOVA, main effect of species: $F_{6,352}$ =365.9, $P<2\times10^{-16}$; main effect of sex: $F_{1,352}$ = 5.3, P=0.02; interaction: $F_{6,352}$ =4.1, P=0.0005). However, there was no evidence that species-level average weights alter median 1hr nest scores (Fig. 4A, phylogenetic generalized least squares, median 1hr nesting score by average weight: coefficient= -0.01, SE=0.06, t=-0.15, P=0.89). Likewise, when we divided the animals by species and sex, we found no correlation between weight and median nest score at 1h (Supplemental Table S4, Spearman's rank correlations, Bonferroni-corrected P >0.05) or maximum overnight nest score (Supplemental Table S4, Spearman's rank correlations, Bonferroni-corrected P >0.05) within any of the species-sex groups. Thus, while average weight varied almost three-fold among species, weight was not associated with nesting behaviour in our assay.

Association between environment and nest construction

We next asked whether there was an association between performance in the nesting assay and several additional environmental covariates, including latitude and average winter temperature of origin, burrow construction, and mating system. Neither latitude nor average winter temperatures were significantly associated with median 1hr scores in these species (Table 2, Fig. 4B; phylogenetic generalized least squares, median 1hr nest score by latitude: coefficient =-0.05, SE= 0.08, t=-0.71, P=0.51; median 1hr nest score by average winter temperature: coefficient = 0.03, SE= 0.05, t=0.69, p=0.52). Moreover, nesting latency does not appear to be influenced by burrowing behaviour: a model in which short nesting latency is dependent on building complex burrows does not fit the data

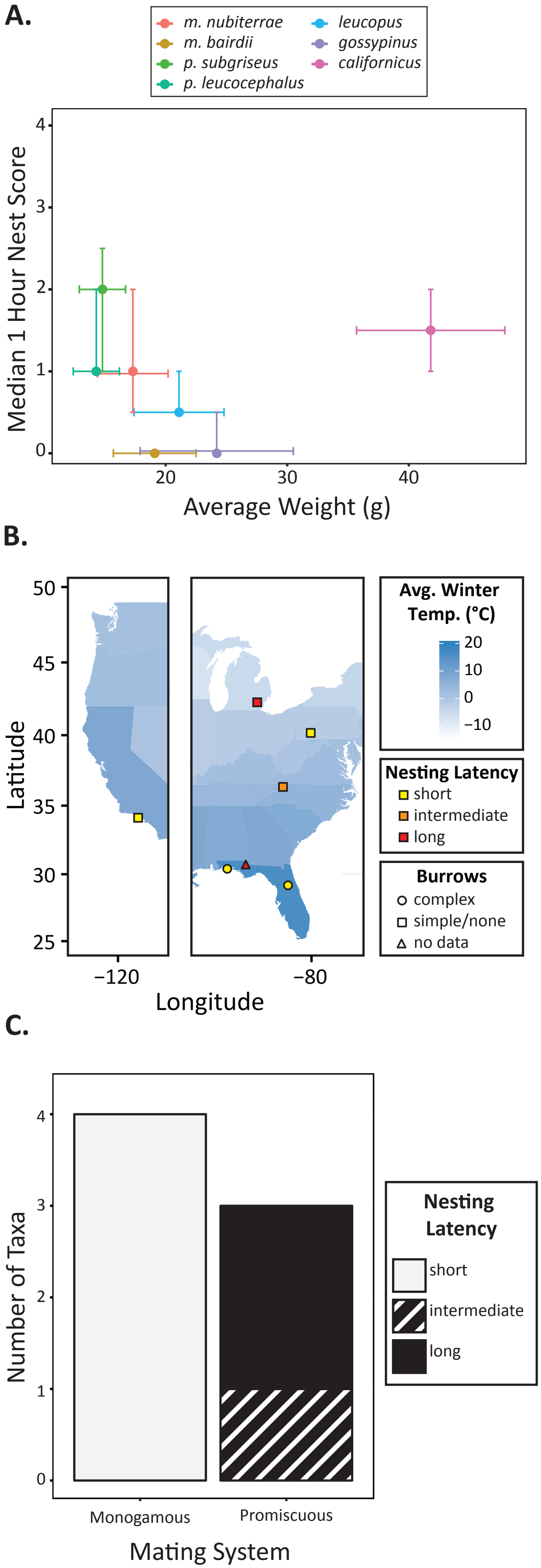


Figure 4: Environmental Factors and Nesting Behaviour

Figure 4: Environmental factors and nesting behaviour. (**A**) Species median 1h nest scores and average weight are plotted with bars indicating interquartile range (nest score) and standard deviation (weight). (**B**) Sites of colony origin (on US map), burrow shape (by symbol), and nesting latency (by colour) are indicated for each taxon (following legend). Map colours represent average winter temperatures by state (Arguez et al., 2010). Nesting latency category (short, intermediate, long) was determined by the significant species groups depicted in Fig. 2B. (**C**) Association between a taxon's mating system and nesting latency.

significantly better than a model where the two traits are independent (Table 2, Fig. 4B; Pagel's binary character correlation test: AIC (independent model) = 21.97, AIC (dependent model) = 23.88, likelihood ratio=2.09, *P*=0.35). However, a model in which short nesting latency depends on mating system fits the observed data significantly better than a model assuming the two traits are independent (Table 2, Fig. 4C; Pagel's binary character correlation test: AIC (independent model)= 26.42, AIC (dependent model)=21.21, likelihood ratio = 9.21, *P*=0.01). With the caveat that the sample size for comparisons among taxa is small, these data suggest that mating system is correlated with nesting latency but the other abiotic environmental factors we examined are not.

DISCUSSION

Nesting is important for survival in rodents, but it is not clear how this behaviour varies among species or which evolutionary pressures drive these changes. Here we designed a novel high-throughput phenotyping paradigm to evaluate variation in both nest structure and the timing of nesting behaviour in closely related species of deer mice. We found that *Peromyscus* mice are generally able to construct dome-shaped nests, but vary strikingly in their latency to do so. Because nesting latency is not simply correlated with phylogeny, this raises the possibility that natural selection may contribute to inter-taxon variation. When we tested for correlations between latency to nest and several abiotic and biotic variables, we found that mating system, but surprisingly not climate or body size, is correlated with nesting behaviour.

Nesting has been well studied in laboratory models (e.g. Lisk et al., 1969; Lynch, 1980). However, the majority of these nesting experiments, including some studies in *Peromyscus*, measure the amount of nesting material that an animal pulls into its cage or the final nest structure achieved over a 24-hour period (Hartung et al., 1979; King et al., 1964; Layne, 1969; Lynch & Hegmann, 1973). By contrast, we focus on both the timing of the behaviour and the final nest structure. By evaluating nests just one hour after the replacement of nesting material, we are able to assess whether the animals differ in their latency to begin nest construction—what might be considered a baseline motivation to nest. This is complemented by a second measurement at the more permissive overnight time point, which allows us to evaluate whether animals vary in their overall ability to build three-dimensional structures. This novel phenotyping paradigm therefore allows us to distinguish between animals that differ in their motivation to construct nests of similar shape from those that differ in their ability to construct nests.

Using this approach, we find that even closely-related *Peromyscus* species vary dramatically in their latency to begin nesting, while variation in final nest structure is much more modest. This is in contrast to studies of nesting in birds and insects, where the structures of complete, species-typical nests are highly variable (Collias, 1964; Healy et al., 2008; Knerer et al., 2012; Price & Griffith, 2017; Schmidt, 1964), or even burrow construction in *Peromyscus*, where species excavate cavities that significantly differ in size and shape (Hu & Hoekstra, 2017; Weber & Hoekstra, 2009). The relative conservation of nest structure implies that the ability to produce dome-shaped nests is important for most animals in the genus. However, variation in latency to begin nesting suggests that prioritization of the behaviour varies among taxa. These patterns also imply that variation in nesting in these mice is likely due to altered motivation rather than changes in stereotyped motor patterns, morphology, or target nest structure.

All animals were acclimated to and tested in a common environment, specifically at 22°C, which is below the preferred temperatures (Ogilvie & Stinson, 1966) and thermoneutral zones (Layne & Dolan, 1975; Glaser & Lustick, 1975; Hayward 1965) of many *Peromyscus* species. While these taxa may differ in their behavioural response to this thermal environment due to differences in basal metabolic rate or thermoneutral zone, we found no evidence of a correlation between nesting scores at either time point and weight, a trait strongly related to both metabolic parameters in *Peromyscus* (Hayward, 1965; Hill, 1983). It is worth noting that the positive relationship between body weight and nest weight observed in previous studies of rodent nesting (King et al., 1964; Lynch, 1992; Wolfe, 1970) might be at least partially explained by larger animals requiring more material to build equivalently shaped structures. By focusing on the structure of the nest rather than the weight of nesting material used to construct it, we minimize this confounding factor.

Body size aside, it is reasonable to hypothesize that climate could alter this thermoregulatory behaviour. Other studies have suggested that climate (King et al., 1964; Lynch, 1992) and microclimate (Wolfe, 1970) alter the amount of nesting material used by rodents in natural populations. However, we find only modest variation in final nest shape and no evidence for a relationship between nesting latency and average winter temperatures or latitude of origin, which is frequently used as a proxy for temperature. Nor do we find evidence for an association between nesting latency and the construction of elaborate burrows, which function as microclimates and buffer the animals from changes in ambient temperature (Hayward, 1965; Sealander, 1952; Weber & Hoekstra, 2009). Although these colonies have experienced reduced selective pressure while bred in laboratory settings, there is still no evidence of an association between climate/microclimate and latency when we consider only the colonies founded within the past 10 years (Table 1; *gossypinus*, *polionotus leucocephalus*, and *maniculatus nubiterrae*). Given that most of the variation we observe takes the form of prioritization differences rather than changes in nest size or shape, it may be especially necessary to consider the broader behavioural repertoire of these taxa, including biotic factors that could contribute to the motivation to nest.

While we do not observe a simple relationship between nesting behaviour and any of the abiotic factors we examined, we do find an intriguing correlation between social environment and nesting latency. With the caveats that our sample size is relatively low and that the classification of species as monogamous or promiscuous relies on incomplete evidence, our results suggest that mating system and nesting latency are not independent, with all putatively monogamous species having short latencies to nest. It is possible that this reflects a tendency to invest in a home territory that is more beneficial for monogamous animals than for promiscuous ones (Gaulin & FitzGerald, 1988), or that selection for increased paternal care, a hallmark of monogamous mating systems (Kleiman, 1977), might result in increased motivation to nest even in virgin animals. This potential relationship between social behaviour and the prioritization of nesting behaviour underscores the importance of considering both biotic and abiotic environment when investigating the causes of behavioural evolution.

CONCLUSION

Measurement of extended phenotypes such as nests allows us to study how behaviours evolve within and between species. Here we showed that the ability to nest is relatively conserved in the genus *Peromyscus*, but latency to begin nest construction is highly variable, even between sister species. This suggests that evolution of nesting behaviour in these animals is characterized by differences in the prioritization of an otherwise conserved behavioural pattern. Intriguingly, while abiotic environment cannot explain these species differences in nesting behaviour, we find a correlation between latency to nest and mating system, with monogamous species prioritizing nesting. Finally, as the innate differences in nesting behaviour in *Peromyscus* appear to be largely changes in the motivation to nest, future studies in this system may elucidate genetic and neurobiological mechanisms that lead to differences in motivation to engage in particular behaviours, a topic with implications far beyond nesting behaviour.

ACKNOWLEDGEMENTS

We would like to thank Harvard undergraduates A. Bielawski, M. Charifson, E. D'Agostino, M. Noriega, and J. Rhodes for blinded scoring of behavioural assays. N. Bedford and C. Hu assisted with the figures. L. Revell and O. Lapiedra provided advice on comparative methods. The Hoekstra lab provided helpful feedback on the experiments and figures, and N. Bedford, E. Hager, D. Haig, C. Hu, B. König, M. Noriega, K. Pritchett-Corning, and K. Turner commented on the manuscript. We would like to thank the Harvard OAR staff for their assistance with animal husbandry.

Funding

CLL was supported by a Morris E. Zuckerman Fellowship, a Smith Family Graduate Science and Engineering Fellowship, and the Harvard Molecules, Cells, and Organisms PhD Program. CLL received a Harvard Mind, Brain, and Behaviour Student Award. HEH is an Investigator of the Howard Hughes Medical Institute. This work was supported, in part, by a National Science Foundation Doctoral Dissertation Improvement Grant (DDIG).

REFERENCES

Arguez, A., Durre, I., Applequist, S., Squires, M., Vose, R., Yin, X., & Bilotta, R. (2010). NOAA's U.S. Climate Normals (1981-2010). (Publication no. DOI:10.7289/V5PN93JP). Retrieved 10/18/16, from NOAA National Centers for Environmental Information.

Bedford, N. L., & Hoekstra, H. E. (2015). *Peromyscus* mice as a model for studying natural variation. *Elife, 4*, e06813.

Blair, W. F. (1950). Ecological factors in speciation of *Peromyscus*. *Evolution*, *4*(3), 253-275.

Blair, W. F. (1951). *Population structure, social behavior, and environmental relations in a natural population of the beach mouse (Peromyscus polionotus leucocephalus)*: University of Michigan Laboratory of Vertebrate Biology Contributions, 48, 1-47.

Bradley, R. D., Durish, N. D., Rogers, D. S., Miller, J. R., Engstrom, M. D., & Kilpatrick, C. W. (2007). Toward a molecular phylogeny for *Peromyscus*: evidence from mitochondrial cytochrome-*b* sequences. *Journal of Mammalogy, 88*(5), 1146-1159.

Bult, A., & Lynch, C. B. (1997). Nesting and fitness: lifetime reproductive success in house mice bidirectionally selected for thermoregulatory nest-building behavior. *Behavior Genetics, 27*(3), 231-240.

Clark, F. H. (1936). Geotropic behavior on a sloping plane of arboreal and non-arboreal races of mice of the genus *Peromyscus*. *Journal of Mammalogy, 17*(1), 44-47.

Collias, N. E. (1964). The evolution of nests and nest-building in birds. *American Zoologist, 4*(2), 175-190.

Crook, J. H. (1963). A comparative analysis of nest structure in the weaver birds (*Ploceinae*) *Ibis, 105*(2), 238-262.

Crossman, C. A., Rohwer, V. G., & Martin, P. R. (2011). Variation in the structure of bird nests between northern Manitoba and southeastern Ontario. *PLoS ONE, 6*(4), e19086.

Dawkins, R. (1982). *The Extended Phenotype: The Long Reach of the Gene*: Oxford University Press.

Dawson, W. D., Lake, C. E., & Schumpert, S. S. (1988). Inheritance of burrow building in *Peromyscus*. *Behavior Genetics, 18*(3), 371-382.

Dewey, M. J., & Dawson, W. D. (2001). Deer mice: the *Drosophila* of North American mammalogy. *Genesis, 29*(3), 105-109.

Dewsbury, D. A., Baumgardner, D. J., Evans, R. L., & Webster, D. G. (1980). Sexual dimorphism for body mass in 13 taxa of muroid rodents under laboratory conditions. *Journal of Mammalogy, 61*(1), 146-149.

Dooley, J. L., & Dueser, R. D. (1990). An experimental examination of nest-site segregation by two *Peromyscus* species. *Ecology, 71*(2), 788-796.

Gaulin, S., & FitzGerald, R. W. (1988). Home-range size as a predictor of mating systems in *Microtus*. *Journal of Mammalogy, 69*(2), 311-319.

Grafen, A. (1989). The phylogenetic regression. *Philosophical Transactions of the Royal Society of London. B , 326*(1233), 119-157.

Glaser, H. & Lustick, S. (1975). Energetics and nesting behavior of the northern white-footed mouse, *Peromyscus leucopus noveboracensis*. *Physiological Zoology, 48*(2), 105-113.

Hansell, M. H. (1984). *Animal Architecture and Building Behaviour*: Longman Press.

Hansell, M. H. (2005). *Animal Architecture*: Oxford University Press.

Hartung, T. G., & Dewsbury, D. A. (1979). Nest-building behavior in seven species of muroid rodents. *Behavioral and Neural Biology, 27*(4), 532-539.

Hayward, J. S. (1965). Microclimate temperature and its adaptive significance in six geographic races of *Peromyscus*. *Canadian Journal of Zoology, 43*(2), 341-350.

Healy, S., Walsh, P., & Hansell, M. (2008). Nest building by birds. *Current Biology, 18*(7), R271-R273.

Hess, S. E., Rohr, S., Dufour, B. D., Gaskill, B. N., Pajor, E. A., & Garner, J. P. (2008). Home improvement: C57BL/6J mice given more naturalistic nesting materials build better nests. *Journal of the American Association for Laboratory Animal Science, 47*(6), 25-31.

Hill, R. W. (1983). Thermal physiology and energetics of *Peromyscus*: Ontogeny, body temperature, metabolism, insulation, and microclimatology. *Journal of Mammalogy, 64*(1): 19-37.

Hu, C. K., & Hoekstra, H. E. (2017). *Peromyscus* burrowing: a model system for behavioral evolution. *Seminars in Cell & Developmental Biology, 61*, 107-114.

Ivey, R. D. (1949). Life history notes on three mice from the Florida east coast. *Journal of Mammalogy, 30*(2), 157-162.

Jašarević, E., Bailey, D. H., Crossland, J. P., Dawson, W. D., Szalai, G., Ellersieck, M. R., Rosenfeld, C. S., & Geary, D. C. (2013). Evolution of monogamy, paternal investment, and female life history in *Peromyscus*. *Journal of Comparative Psychology, 127*(1), 91-102.

King, J. A., Maas, D., & Weisman, R. (1964). Geographic variation in nest size among species of *Peromyscus*. *Evolution*, *18*(2),230-234.

Kleiman, D. G. (1977). Monogamy in mammals. *The Quarterly Review of Biology, 52*(1), 39-69.

Klein, H. G., & Layne, J. N. (1978). Nesting behavior in four species of mice. *Journal of Mammalogy, 59*(1), 103-108.

Knerer, G., & Atwood, C. E. (2012). Nest architecture as an aid in Halictine taxonomy (Hymenoptera: Halictidae). *The Canadian Entomologist, 98*(12), 1337-1339.

Lackey, J. A., Huckaby, D. G., & Ormiston, B. G. (1985). *Peromyscus leucopus*. *Mammalian Species Archive, 247*, 1-10.

Layne, J. N. (1969). Nest-building behavior in three species of deer mice, *Peromyscus*. *Behaviour, 35*(3), 288-302.

Layne, J. N. & Dolan, P. G. (1975). Thermoregulation, metabolism, and water economy in the golden mouse (*Ochroyomys nuttalli*). *Comparative Biochemistry and Physiology Part A: Physiology*, *52*(1), 153-163.

Lisk, R. D., Pretlow , R. A., & Friedman, S. M. (1969). Hormonal stimulation necessary for elicitation of maternal nest-building in the mouse (*Mus musculus*). *Animal Behaviour, 17*(4), 730-737.

Lynch, C. B. (1980). Response to divergent selection for nesting behavior in *Mus musculus*. *Genetics, 96*(3), 757-765.

Lynch, C. B. (1992). Clinal variation in cold adaptation in *Mus domesticus*: verification of predictions from laboratory populations. *The American Naturalist, 139*(6), 1219-1236.

Lynch, C. B., & Hegmann, J. P. (1973). Genetic differences influencing behavioral temperature regulation in small mammals. II. Genotype-environment interactions. *Behavior Genetics, 3*(2), 145-154.

Lynch, C. B., & Possidente, B. P. (1978). Relationships of maternal nesting to thermoregulatory nesting in house mice (*Mus musculus*) at warm and cold temperatures. *Animal Behaviour, 26*, 1136-1143.

M'Closkey, R. T. (1976). Community structure in sympatric rodents. *Ecology, 57*(4), 728-739.

Mainwaring, M. C., Hartley, I. R., Lambrechts, M. M., & Deeming, D. C. (2014). The design and function of birds' nests. *Ecology and Evolution, 4*(20), 3909-3928.

McCarley, W. (1959). A study of the dynamics of a population of *Peromyscus gossypinus* and *P. nuttalli* subjected to the effects of X-irradiation. *American Midland Naturalist*, *61*(2), 447-469.

Merritt, J. F. (1974). Factors influencing the local distribution of *Peromyscus californicus* in Northern California. *Journal of Mammalogy, 55*(1), 102-114.

Merritt, J. F. (1978). *Peromyscus californicus*. *Mammalian Species* (85), 1-6.

Meserve, P. L. (1977). Three-dimensional home ranges of cricetid rodents. *Journal of Mammalogy, 58*(4), 549-558.

Morris, J. G., & Kendeigh, S. C. (1981). Energetics of the prairie deer mouse *Peromyscus maniculatus bairdii*. *American Midland Naturalist*, *105*(2), 368-376.

Ogilvie, D. M., & Stinson, R. H. (1966). Temperature selection in *Peromyscus* and laboratory mice, *Mus musculus*. *Journal of Mammalogy*, *47*(4), 655-660.

Pagel, M. (1994). Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proceedings of the Royal Society of London B,255*(1342), 37-45.

Paradis E., Claude, J. & Strimmer, J. (2004). APE: analyses of phylogenetics and evolution in R language. *Bioinformatics, 20*, 289-290.

Pearson, O. P. (1960). The oxygen consumption and bioenergetics of harvest mice. *Physiological Zoology, 33*(2), 152-160.

Pearson, P. G. (1953). A field study of *Peromyscus* populations in Gulf Hammock, Florida. *Ecology, 34*(1), 199-207.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & Team, R. C. (2017). nlme: Linear and Nonlinear Mixed Effects Models (Version R package version 3.1-131).

Price, J. J., & Griffith, S. C. (2017). Open cup nests evolved from roofed nests in the early passerines. *Proceedings of the Royal Society B, 284*(1848), online early edition.

Revell, L. J. (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution, 3*, 217-223.

Schmidt, R. S. (1964). *Apticotermes* nests. *American Zoologist, 4*(2), 221-225.

Sealander, J. A. (1952). The relationship of nest protection and huddling to survival of *Peromyscus* at low temperature. *Ecology, 33*(1), 63-71.

Southwick, C. H. (1955). Regulatory mechanisms of house mouse populations: social behavior affecting litter survival. *Ecology, 36*(4), 627-634.

Sumner, F. B. (1926). An analysis of geographic variation in mice of the *Peromyscus polionotus* group from Florida and Alabama. *Journal of Mammalogy, 7*(3), 149-184.

Symonds, M. R. E., & Blomberg, S. P. (2014). A primer on phylogenetic generalised least squares. In L. Z. Garamszegi (Ed.), *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology: Concepts and Practice* (pp. 105-130). Berlin, Heidelberg: Springer Berlin Heidelberg.

Turner, L. M., Young, A. R., Römpler, H., Schöneberg, T., Phelps, S. M., & Hoekstra, H. E. (2010). Monogamy evolves through multiple mechanisms: evidence from V1aR in deer mice. *Molecular Biology and Evolution, 27*(6), 1269-1278.

Vogt, F. D., & Lynch, G. R. (1982). Influence of ambient temperature, nest availability, huddling, and daily torpor on energy expenditure in the white-footed mouse *Peromyscus leucopus*. *Physiological Zoology*, *55*(1), 56-63.

Weber, J. N., & Hoekstra, H. E. (2009). The evolution of burrowing behaviour in deer mice (genus *Peromyscus*). *Animal Behaviour, 77*(3), 603-609.

Winkler, D. W., & Sheldon, F. H. (1993). Evolution of nest construction in swallows (*Hirundinidae*): a molecular phylogenetic perspective. *Proceedings of the National Academy of Sciences, 90*(12), 5705-5707.

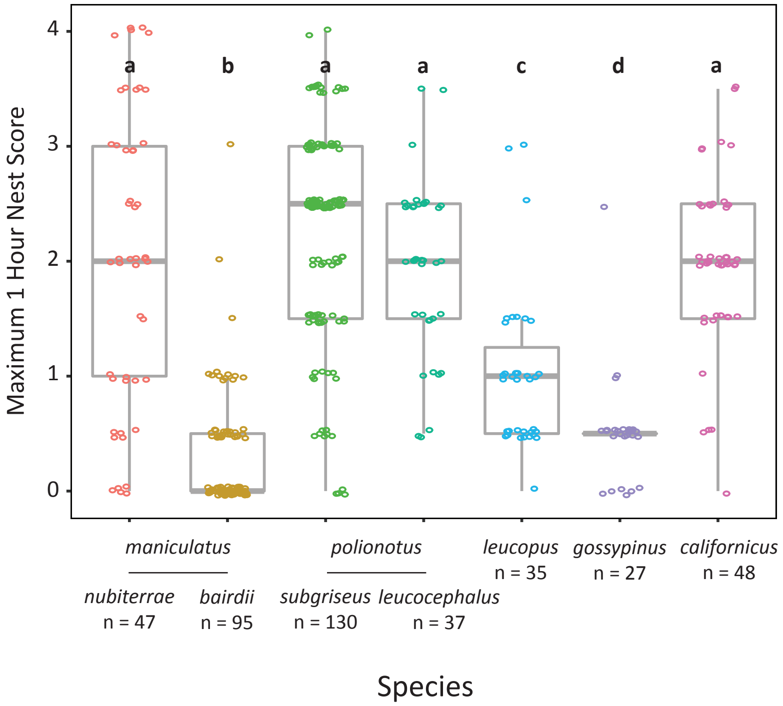
Wolfe, J. L. (1970). Experiments on nest-building behaviour in *Peromyscus* (Rodentia: Cricetinae). *Animal Behaviour, 18*, 613-615.

Wolfe, J. L., & Linzey, A. V. (1977). *Peromyscus gossypinus*. *Mammalian Species,* (70), 1-5.

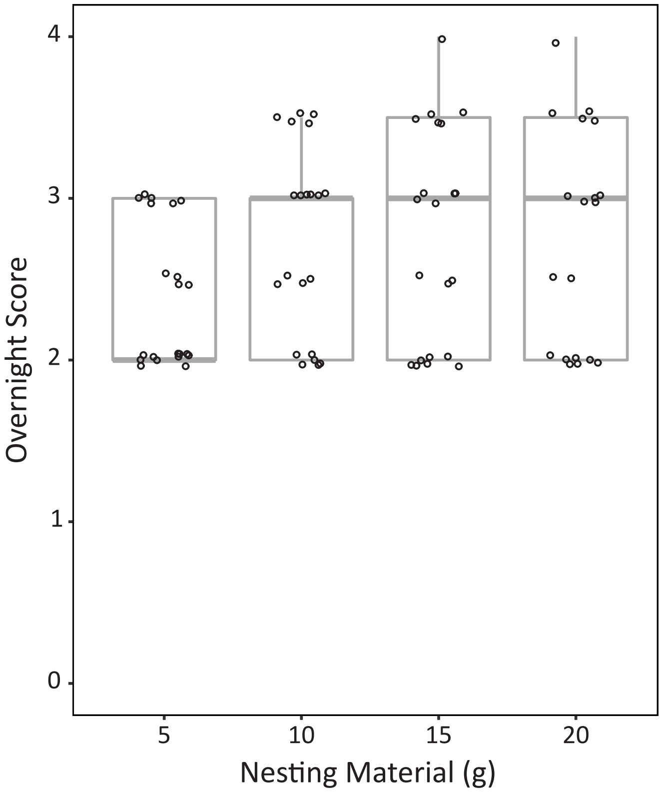
Wolff, J. O., & Cicirello, D. M. (1991). Comparative paternal and infanticidal behavior of sympatric white-footed mice (*Peromyscus leucopus noveboracensis*) and deermice (*P. maniculatus nubiterrae*). *Behavioral Ecology, 2*(1), 38-45.

Wolff, J. O., & Durr, D. S. (1986). Winter nesting behavior of *Peromyscus leucopus* and *Peromyscus maniculatus*. *Journal of Mammalogy, 67*(2), 409-412.

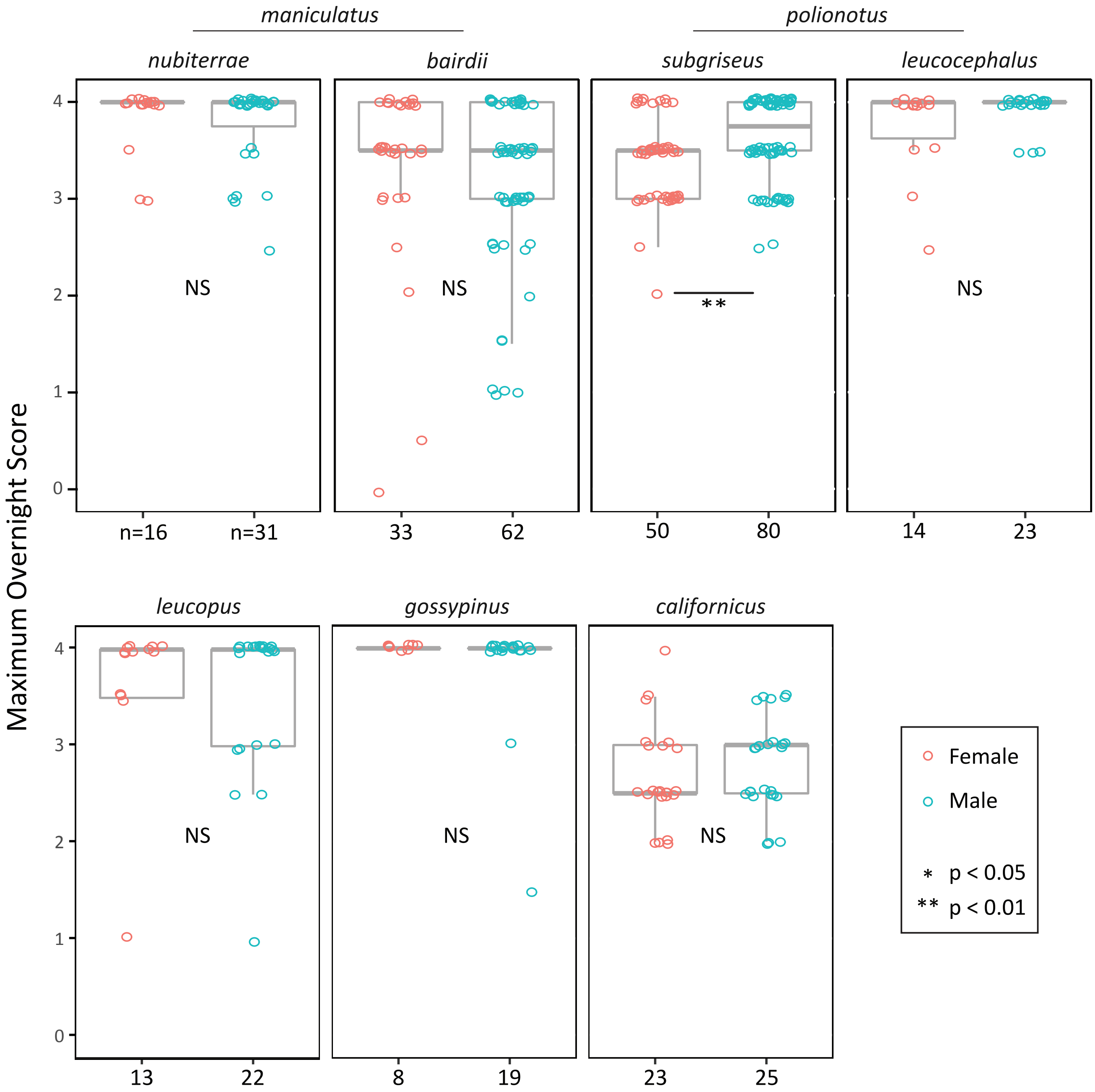
Wolff, J. O., & Hurlbutt, B. (1982). Day refuges of *Peromyscus leucopus* and *Peromyscus maniculatus*. *Journal of Mammalogy, 63*(4), 666-668.



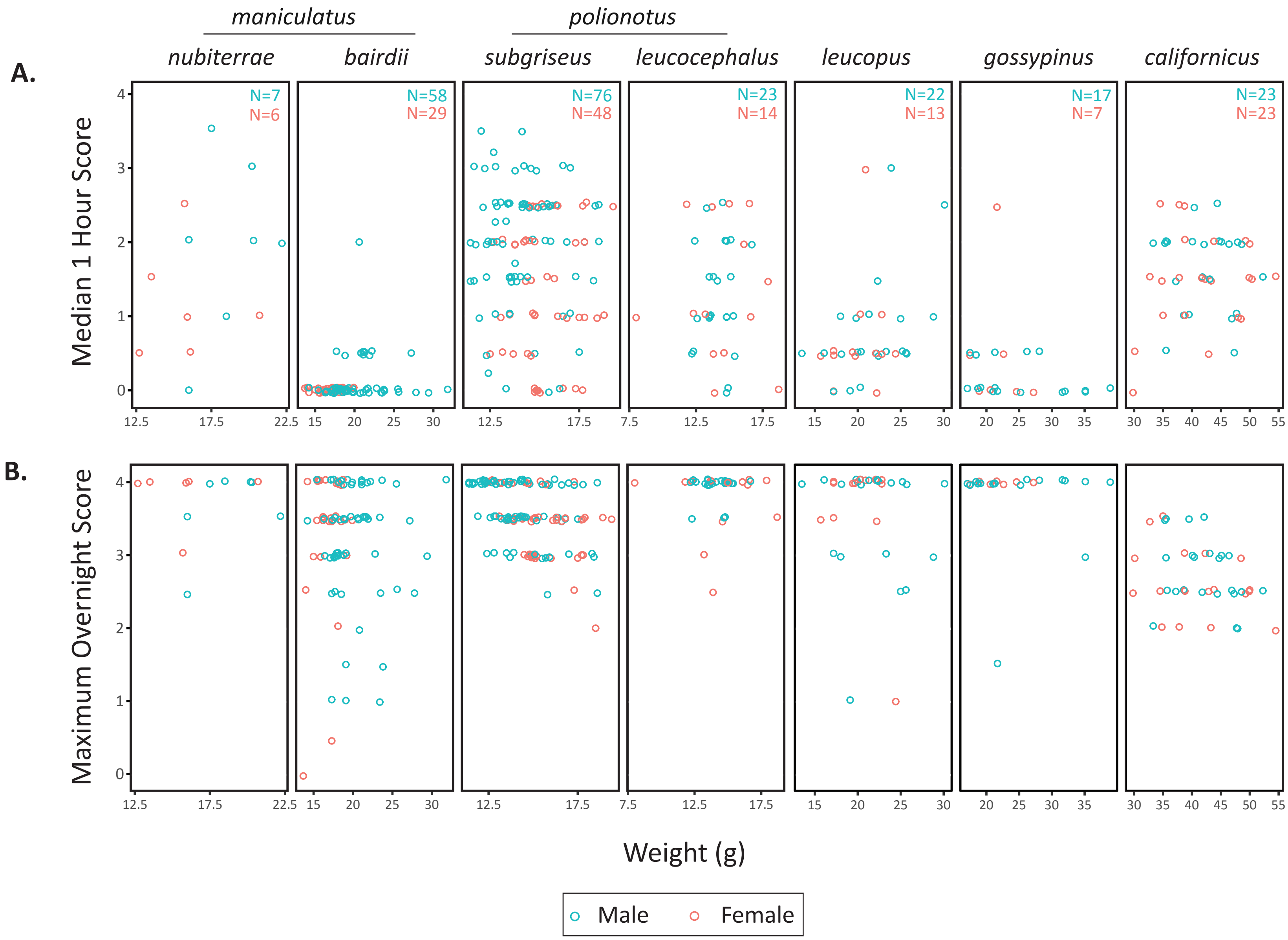
Supplemental Figure S1: Maximum 1 hour nest scores. Maximum nest score achieved 1h after receiving new nesting material over the 3 trial days. Letters indicate species groups that did not significantly differ from one another, while all other pairwise comparisons were significant (Wilcoxon rank-sum test, Bonferroni-corrected *P*<0.05), largely consistent with median scores (Fig. 2B, Supplemental Table S2), with the exception of a significant difference in maximum scores between *P. gossypinus* and *P. m. bairdii* animals (Wilcoxon rank-sum test, Bonferroni-corrected *P*=0.049). Sample sizes provided below.



Supplemental Figure S2: Effect of increasing nesting material in *P. californicus*. Adult animals (*N*=21) were given increasing amounts of nesting material (5g, 10g, 15g and 20g) on 4 sequential days. Higher amounts of nesting material increased overnight nest scores (Friedman repeated measures test, *P*=0.003).



Supplemental Figure S3: Sex differences in maximum overnight nest scores. Evidence for a significant sex difference overnight nest score occurred in only one taxon: in *P. polionotus subgriseus*, males built higher scoring nests at the overnight time point than females (Wilcoxon rank-sum test, Bonferroni-corrected *P*=0.008). Sample sizes are provided below. NS = not significant.



Supplemental Figure S4: Effect of body weight on nesting behaviour by species and sex. (A) Median nest scores at 1h were not correlated with weight within any species-sex group (Spearman correlation, Bonferroni-corrected *P*>0.05). (B) Overnight maximum nest scores were not correlated with weight within any species-sex group (Spearman correlation, Bonferroni-corrected *P*>0.05). Note different x-axis for each species as indicated. Sample sizes are provided by sex and were the same for both time points.

Supplemental Table S1: Detailed Nest Scoring Criteria

Score	***Score Description***	***Shredding***	***Nest Site[1]***	***Walls[2]***	***Overhead Cover***
0	**No Manipulation**	None	-	-	-
0.5	**Minor Shredding**	Minor: ≤top of 1 nestlet	-	-	-
1	**Extensive Shredding**	Extensive: > top of 1 nestlet	No: neither a nor b is true	No	-
1.5	**Ambiguous Nest Site**	Extensive	Unclear: either a OR b is true	No	-
2	**Platform Nest**	-	Yes: both a and b are true	No: < ½ sphere height for < ½ circumference	-
2.5	**Partial Cup Nest**	-	Yes	Partial: < ½ sphere height for > ½ circumference OR ≥½ sphere height for < ½ of circumference	-
3	**Cup Nest**	-	Yes	Yes: ≥½ sphere height for ≥ ½ circumference	no overhead cover
3.5	**Partial Dome Nest**	-	Yes	Yes	Partial: <50% of the sphere is covered or there are multiple entrance holes
4	**Full Dome Nest**	-	Yes	Yes	Yes: ≥50% of the sphere is covered and there is at most one entrance hole

1. A nest site is defined according to two criteria: (**a**) there is a contiguous concentration of nestlet around a central point consisting of ≥90% of any material the animal has shredded, and (**b**) the shape of the putative nest site is defined by the shredded material (>50% of the nestlet at the nest site is shredded).
2. To evaluate walls, imagine that the nest cavity is filled by a sphere, *sensu* (Hess et al., 2008). The walls are compared to the height of the sphere within the cavity, and the proportion of the circumference of the sphere that is surrounded by walls is noted.

Supplemental Table S2: Pairwise species comparisons of 1hr median and overnight maximum scores

1hr Median Scores

Overnight Maximum Scores

Species	***maniculatus nubiterrae***	***maniculatus bairdii***	***polionotus subgriseus***	***polionotus leucoceph.***	***leucopus***	***gossypinus***	***californicus (5g)***	***californicus (20g)***
N:	**47**	**95**	**130**	**37**	**35**	**27**	**48**	**23**
man. nub.		*W*=3224.5, **P=9.3x10^{-5}**	*W*=4124.5, **P=0.0026**	*W*=819.5, *P*=1	*W*=905, *P*=1	*W*=540, *P*=1	*W*=2102, **P=8.9x10^{-13}**	*W*=974.5 **P=7.6x10^{-8}**
man. baird.	*W*=4046, **P<2.2x10^{-16}**		*W*=5306.5, *P*=1	*W*=846, **P=1.8x10^{-05}**	*W*=1133, *P*=0.073	*W*=545, **P=2.8x10^{-5}**	*W*=3473, **P=3.8x10^{-6}**	*W*=1553.5 **P=0.027**
pol. sub.	*W*=2565, *P*=1	*W*=734, **P<2.2x10^{-16}**		*W*=1353.5, **P=0.00025**	*W*=1778, *P*=0.72	*W*=866, **P=00017**	*W*=5435.5, **P=8.6x10^{-14}**	*W*=2428.5 **P=1.3x10^{-5}**
pol. leuc.	*W*=924.5, *P*=1	*W*=256.6, **P<2.2x10^{-16}**	*W*=3088.5, *P*=0.16		*W*=746.5, *P*=1	*W*=447.5, *P*=1	*W*=1693, **P=2.9x10^{-12}**	*W*=791 **P=3.9x10^{-8}**
leucop.	*W*=1159.5, **P=0.027**	*W*=375, **P=1.7x10^{-15}**	*W*=3589.5, **P=2.5x10^{-6}**	*W*=929.5, **P=0.023**		*W*=362, *P*=0.59	*W*=1438.5, **P=2.9x10^{-7}**	*W*=666 **P=0.00024**
gossy.	*W*=1095, **P=2.3x10^{-6}**	*W*=1036, *P*=0.29	*W*=3176, **P=4.6x10^{-10}**	*W*=881.5, **P=1.8x10^{-6}**	*W*=763.5, **P=0.0002**		*W*=1221, **P=1.2x10^{-9}**	*W*=576 **P=4.9x10^{-7}**
califor.	*W*=990.5, *P*=1	*W*=103, **P<2.2x10^{-16}**	*W*=3492, *P*=1	*W*=678.5, *P*=1	*W*=281, **P=3.1x10^{-6}**	*W*=74.5, **P=2.4x10^{-9}**		

The results of pairwise Wilcoxon rank-sum tests for species differences in median 1hr scores (below diagonal) or maximum overnight scores (above diagonal). For each comparison, test statistics (W) and Bonferroni-corrected p-values are reported; significant results (p<0.05) are in bold.

Supplemental Table S3: Sex differences in nest scores

Taxon	Males	Females	1hr Median Score	Maximum Overnight Score
maniculatus nubiterrae	31	16	***W*=119.5, *P*=0.03**	*W*=266, *P*=1
maniculatus bairdii	62	33	*W*=858, *P*=0.11	*W*=1173, *P*=1
polionotus subgriseus	80	50	***W*=1262.5, *P*=0.002**	***W*=1361, *P*=0.008**
polionotus leucocephalus	23	14	*W*=163.5, *P*=1	*W*=133, *P*=1
leucopus	22	13	*W*=125.5, *P*=1	*W*=152, *P*=1
gossypinus	19	8	*W*=87, *P*=1	*W*=84, *P*=1
californicus	25	23	*W*=219.5, *P*=1	*W*=238.5, *P*=1

The results of pairwise Wilcoxon rank-sum tests for sex differences in 1hr median nest scores or maximum overnight nest scores. Sample sizes, test statistics (W), and Bonferroni-corrected p-values are reported; significant results (P<0.05) are in bold.

Supplemental Table S4: Spearman correlations between weight and nest scores within species/sex groups

	Weight vs. 1hr Median Score		**Weight vs. Maximum Overnight Score**	
Taxon	**Males**	**Females**	**Males**	**Females**
maniculatus nubiterrae	r_s=0.24, *N*=7, *P*=1	r_s=-0.09, *N*=6, *P*=1	r_s=0.39, *N*=7, *P*=1	r_s=0.13, *N*=6, *P*=1
maniculatus bairdii	r_s=0.22, *N*=58, *P*=1	*N*=29, see below	r_s=0.07, *N*=58, *P*=1	r_s=0.13, *N*=29, *P*=1
polionotus subgriseus	r_s=0.01, *N*=76, *P*=1	r_s=0.04, *N*=48, *P*=1	r_s=-0.32, *N*=76, *P*=0.08	r_s=-0.31, *N*=48, *P*=0.48
polionotus leucocephalus	r_s=0.03, *N*=23, *P*=1	r_s=-0.10, *N*=14, *P*=1	r_s=0.11, *N*=23, *P*=1	r_s=-0.06, *N*=14, *P*=1
leucopus	r_s=0.41, *N*=22, *P*=0.82	r_s=0.22, *N*=13, *P*=1	r_s=-0.08, *N*=22, *P*=1	r_s=-0.04, *N*=13, *P*=1
gossypinus	r_s=-0.25, *N*=17, *P*=1	r_s=-0.24, *N*=7, *P*=1	r_s=-0.23, *N*=17, *P*=1	*N*=7, see below
californicus	r_s=-0.08, *N*=23, *P*=1	r_s=0.14, *N*=23, *P*=1	r_s=-0.36, *N*=23, *P*=1	r_s=-0.29, *N*=23, *P*=1

Sample sizes, Spearman correlation coefficient ($\mathrm{r_s}$), and Bonferroni-corrected p-values are reported. Sample sizes are smaller than for other tests due to missing weight data. We were unable to perform correlations between 1hr scores and weight within *maniculatus bairdii* females because all 29 animals received a median score of 0 at 1 hour. Similarly, all female *gossypinus* animals produced maximum scores of 4 at the overnight time point.