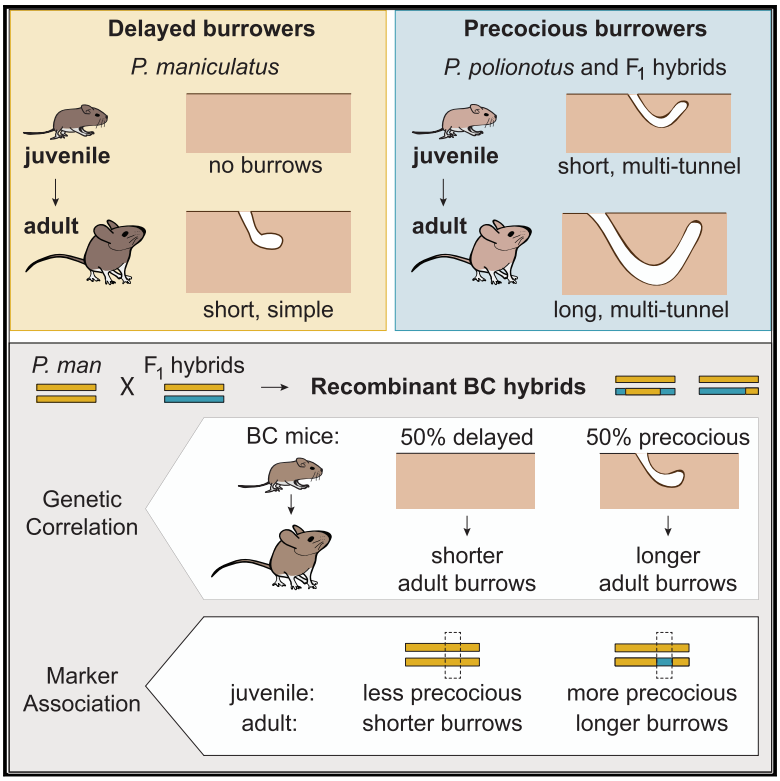
Report

Current Biology

Evolution and Genetics of Precocious Burrowing Behavior in *Peromyscus* Mice

Graphical Abstract



Highlights

- *P. polionotus* build miniaturized, adult-like burrows as early as 19 days old
- Juvenile *P. polionotus* construct burrows precociously compared to *P. maniculatus*
- Cross-fostering does not alter species-specific burrowing behavior
- A *P. polionotus* allele increases adult tunnel length and early onset of burrowing

Authors

Hillery C. Metz, Nicole L. Bedford, Yangshu Linda Pan, Hopi E. Hoekstra

Correspondence

hoekstra@oeb.harvard.edu

In Brief

Metz et al. find that oldfield mice, a species that digs long, complex burrows, also digs burrows earlier than its sister species. In hybrids, an allele linked to adult tunnel length also affects the timing of first burrow construction, suggesting that this genetic region influences different aspects of the same behavior across life stages.

Metz et al., 2017, Current Biology *27*, 3837–3845
December 18, 2017
https://doi.org/10.1016/j.cub.2017.10.061

Current Biology
Report

Evolution and Genetics of Precocious Burrowing Behavior in *Peromyscus* Mice

Hillery C. Metz,[1,2,3] Nicole L. Bedford,[1,2,3] Yangshu Linda Pan,[1,2] and Hopi E. Hoekstra[1,2,4,*]
[1]Department of Organismic & Evolutionary Biology, Department of Molecular & Cellular Biology, Center for Brain Science, and the Museum of Comparative Zoology, Harvard University, Cambridge, MA 02138, USA
[2]Howard Hughes Medical Institute
[3]These authors contributed equally
[4]Lead Contact
*Correspondence: hoekstra@oeb.harvard.edu
https://doi.org/10.1016/j.cub.2017.10.061

SUMMARY

A central challenge in biology is to understand how innate behaviors evolve between closely related species. One way to elucidate how differences arise is to compare the development of behavior in species with distinct adult traits [1]. Here, we report that *Peromyscus polionotus* is strikingly precocious with regard to burrowing behavior, but not other behaviors, compared to its sister species *P. maniculatus*. In *P. polionotus*, burrows were excavated as early as 17 days of age, whereas *P. maniculatus* did not build burrows until 10 days later. Moreover, the well-known differences in burrow architecture between adults of these species—*P. polionotus* adults excavate long burrows with an escape tunnel, whereas *P. maniculatus* dig short, single-tunnel burrows [2–4]—were intact in juvenile burrowers. To test whether this juvenile behavior is influenced by early-life environment, we reciprocally cross-fostered pups of both species. Fostering did not alter the characteristic burrowing behavior of either species, suggesting that these differences are genetic. In backcross hybrids, we show that precocious burrowing and adult tunnel length are genetically correlated and that a *P. polionotus* allele linked to tunnel length variation in adults is also associated with precocious onset of burrowing in juveniles, suggesting that the same genetic region—either a single gene with pleiotropic effects or linked genes—influences distinct aspects of the same behavior at these two life stages. These results raise the possibility that genetic variants affect behavioral drive (i.e., motivation) to burrow and thereby affect both the developmental timing and adult expression of burrowing behavior.

RESULTS

P. polionotus Construct Burrows Earlier in Life than *P. maniculatus*

To examine the developmental onset of burrow construction in *Peromyscus* mice, we assayed burrowing behavior in juveniles starting at 17 days of age (these mice are typically weaned at postnatal day 24 [P24]). We found striking interspecific differences in both the timing and progression of burrow construction (Figure 1; Table S1). Notably, *P. polionotus* were precocious diggers, constructing complete burrows—defined as excavations with at least two components, an entrance tunnel and a nest chamber—on average 10 days earlier than *P. maniculatus*. The first appearance of a complete burrow was at P17 in *P. polionotus* (1 of 5 mice; Figure 1B), but not until P27 in *P. maniculatus* (3 of 14 mice; Figure 1B), a considerable difference in developmental stage (see Figure S1 for timeline of development). Moreover, *P. polionotus* burrowed at adult-like frequencies from P19 onward, a developmental benchmark *P. maniculatus* did not reach until P27 (Figure 1B; Table S1).

Whereas tunnel length increased with age in both species, reflecting a progression in burrowing ability with growth and development (Figure 1C; analysis of covariance [ANCOVA], $p < 0.0001$), tunnel length varied considerably between species. *P. polionotus* consistently produced significantly longer burrows than *P. maniculatus* (Figure 1C; ANCOVA, $p < 0.0001$; Cohen's $d = 1.79$), consistent with the known differences in adult tunnel length [2–4]. Furthermore, the rate of increase in tunnel length across ontogeny was significantly greater for *P. polionotus* (Figure 1C; ANCOVA, age × species interaction, $p = 0.023$). Thus, both the expression of adult-like burrowing frequency and an increase in excavation length develops more rapidly in *P. polionotus* than in *P. maniculatus*.

In trials in which mice did not construct full burrows, individuals of both species usually excavated shallow cup-shaped cavities (divots) instead. Only three of 97 mice (two P17 *P. polionotus* and one P27 *P. maniculatus*) failed to leave any signs of digging activity. These data suggest that the motor patterns for digging were partly, if not completely, developed in both species by at least P17.

Juveniles Construct Burrows with Miniaturized Adult Architecture

Juveniles from both species produced burrows with architecture typical of adults of their respective species. *P. polionotus* constructed escape tunnels as early as P19, and by P21, their burrows included escape tunnels (4 of 7 mice) as frequently as conspecific adults (6 of 9 mice) (Figure 1D; Fisher's exact test, one-tailed, $p = 0.549$). Likewise, *P. maniculatus* juvenile burrows invariably featured only a single tunnel leading to the nest chamber, always lacking an escape tunnel (Figure 1D). Although

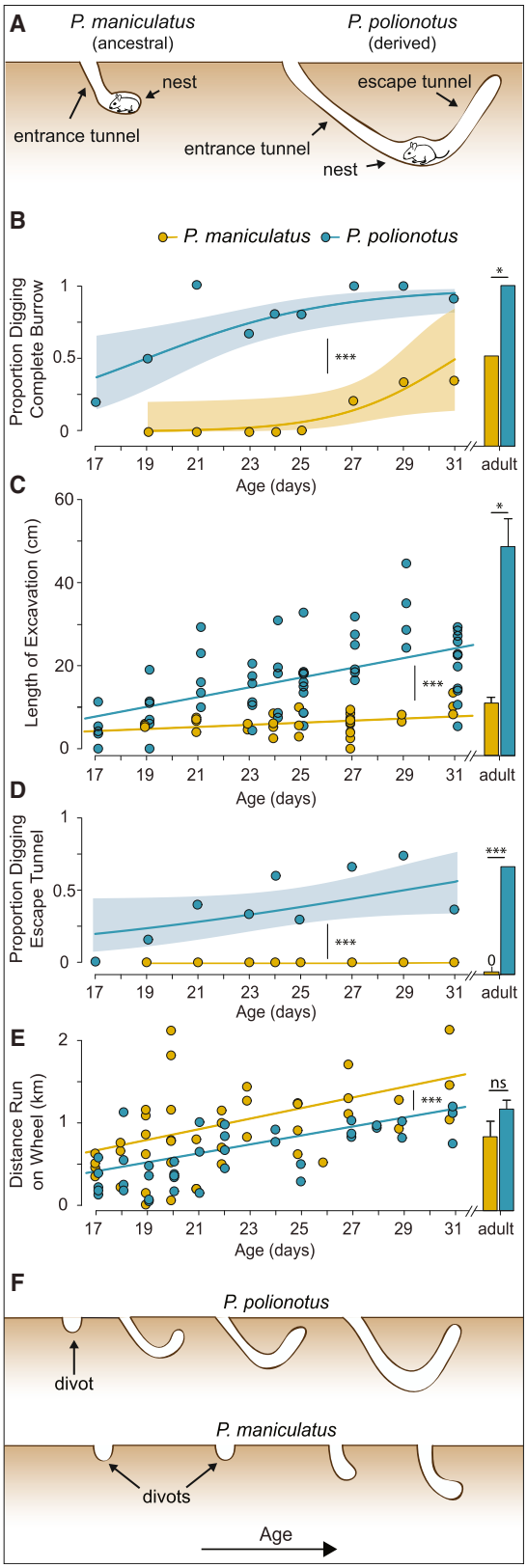


Figure 1. The Ontogeny of Burrow Construction in Two Sister Species of *Peromyscus*

All *P. maniculatus* (yellow) and *P. polionotus* (blue) were naive and were tested only once.

(A) The ancestral burrow architecture, built by *P. maniculatus*, is short (<15 cm) and simple. In contrast, adult *P. polionotus* dig stereotyped burrows with a long entrance tunnel, nest chamber, and escape tunnel (total excavation length ~50 cm).

(B) Proportion of tested mice constructing a complete burrow (i.e., entrance tunnel and nest chamber). Curves and shaded areas represent binary generalized linear smoothers with 95% confidence intervals. Species differences were evaluated by Fisher's exact test (juveniles: see main text for details; adults: *P. maniculatus*, n = 17, and *P. polionotus*, n = 9).

(C) Length of total excavation. Juvenile differences evaluated by ANCOVA (see main text for details). Adult differences between species were evaluated by t test (*P. maniculatus*, n = 17; *P. polionotus*, n = 9). Error bars indicate ±1 SEM.

(D) Proportion of tested mice constructing an escape tunnel. Statistical tests are as in (B).

(E) Distance run on a wheel during a 90-min trial by *P. maniculatus* and *P. polionotus* juveniles (P17–P31; see main text for details) and adults (>P60; *P. maniculatus*, n = 10; *P. polionotus*, n = 10). Statistical tests are as in (C).

(F) Cartoon depiction of data shown in (B)–(D) highlighting the variation in burrow shape over development.

Significance levels are indicated as follows: ns (not significant), $p \geq 0.05$; *$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$. See also Figure S1 and Table S1.

complete with regard to architectural components, juvenile excavations were significantly shorter than those of adults (Figure 1C; t tests, $p < 0.0001$ for both species; *P. maniculatus* Cohen's $d = 1.12$; *P. polionotus* Cohen's $d = 2.01$), thus representing miniature versions of adult burrows.

Precociousness Is Specific to Burrowing Behavior

To evaluate whether precocious burrow construction in *P. polionotus* might be due to advantages in physical rather than behavioral development (e.g., [5]), we examined general measures of morphological and motor development in both species. Two lines of evidence refute this hypothesis. First, *P. polionotus* did not perform better in a second motor activity task: *P. polionotus* juveniles traveled less distance in a 90-min wheel-running assay than *P. maniculatus*. Although total distance run increased with age at a comparable rate in both species (Figure 1E; age × species interaction term, $p = 0.599$), *P. maniculatus* ran significantly greater distances than age-matched *P. polionotus* (ANCOVA, $p < 0.001$). Second, *P. polionotus* are smaller than *P. maniculatus* in both body mass (ANCOVA, $p < 0.0001$) and hindfoot length (ANCOVA, $p < 0.0001$) across development (Figure S1). Likewise, we did not observe heterochrony favoring *P. polionotus* with respect to additional developmental milestones, as *P. maniculatus* reached them earlier in life (Figure S1). Thus, precocious burrowing in *P. polionotus* juveniles reflects a behavioral difference, most likely specific to burrowing, not an advantage in overall activity level, motor ability, or morphological development.

Species-Specific Burrowing Behavior Unaltered by Interspecific Cross-fostering

To disentangle the effects of genetics from environment, we reciprocally cross-fostered pups between the two sister species (Figure 2A). We reasoned that any effects on burrowing behavior resulting from parental environment were likely to be greatest during post-natal development.

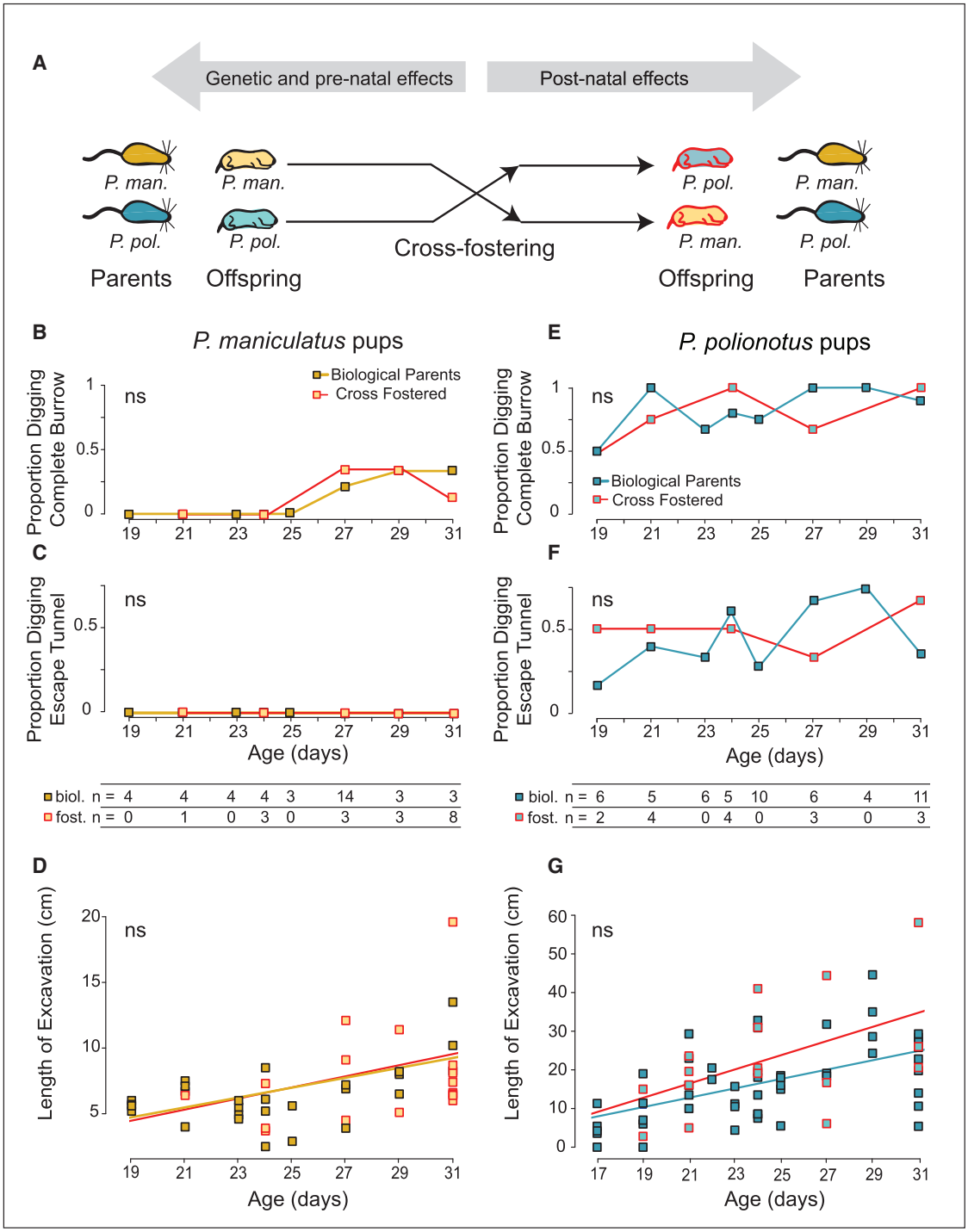


Figure 2. Reciprocal Interspecific Cross-fostering

(A) Schematic of cross-fostering design with *P. maniculatus* (yellow) and *P. polionotus* (blue), with cross-fostered pups highlighted in red.

(B–G) Proportion of mice constructing complete burrows (B and E), proportion of mice building an escape tunnel (C and F), and length of excavations (D and G). Sample sizes for each age and foster group are shown. For (B), (C), (E), and (F), differences between foster treatments were evaluated by Fisher's exact test; for (D) and (G), they were evaluated by ANCOVA (see main text for details). Significance levels are indicated as follows: ns (not significant), $p \geq 0.05$.

In *P. maniculatus*, the developmental onset of burrow building did not differ between cross-fostered and non-fostered animals. Prior to P27, *P. maniculatus* juveniles did not build complete burrows regardless of foster treatment (Figure 2B). After the onset of burrowing, fostered animals constructed burrows no more frequently (4 of 14 mice) than pups reared by their biological

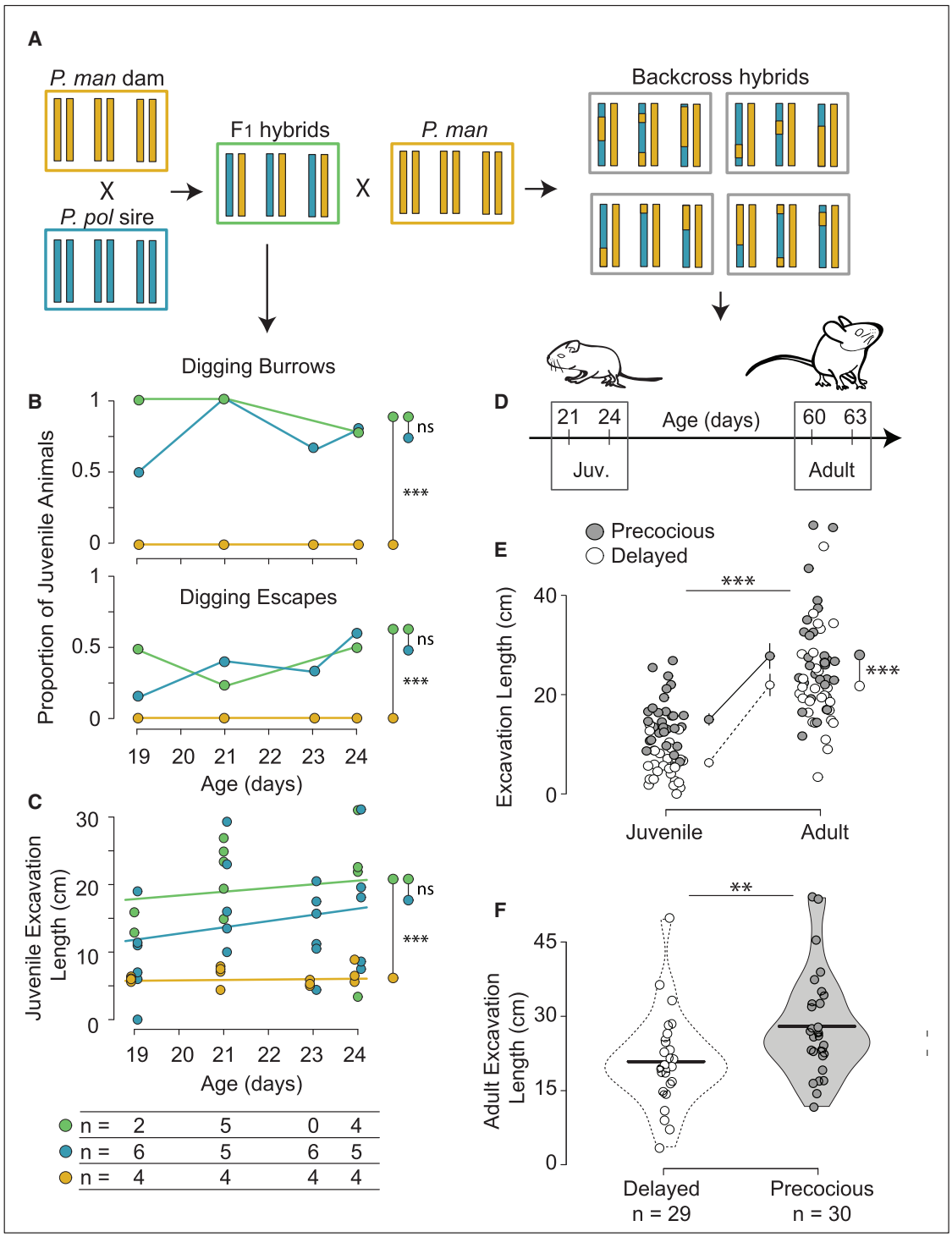


Figure 3. Genetic Dissection of Precocious Burrowing in *P. polionotus* × *P. maniculatus* Hybrids

(A) Schematic of breeding design showing *P. maniculatus* (yellow), *P. polionotus* (blue), first-generation (F_1) hybrids (green), and second-generation backcross (BC) hybrids (gray).

(B) Proportion of juvenile animals digging complete burrows (top) and escape tunnels (bottom); groups were compared using Fisher's exact tests (see main text for details).

(C) Length of excavations in F_1 hybrids compared to *P. maniculatus* and *P. polionotus*; species differences were evaluated by ANCOVA (see main text for details). Sample sizes for each group are shown below.

(D) Timeline of the four behavioral assays completed for each BC hybrid.

(legend continued on next page)

parents (5 of 20 mice) (Figure 2B; Fisher's exact test, one-tailed, p = 0.560). Cross-fostered *P. maniculatus* also did not build escape tunnels (Figure 2C), and the excavations of cross-fostered animals closely matched those of mice raised by their biological parents with regard to length (Figure 2D; ANCOVA, p = 0.485; Cohen's *d* = 0.16).

Likewise, *P. polionotus* raised by heterospecific parents began burrowing at the earliest age tested (P19; Figure 2E), and from P21 onward, nearly all cross-fostered *P. polionotus* excavated burrows (12 of 14 mice; Figure 2E). Burrow structure also did not change with cross-fostering treatment. Cross-fostered *P. polionotus* dug escape tunnels as early in ontogeny (from P19), and as frequently (50%, 8 of 16 mice), as non-fostered juveniles (41%, 22 of 53 mice; Figure 2F, Fisher's exact test, one-tailed, p = 0.813) and conspecific adults (67%, 6 of 9 mice; Fisher's exact test, one-tailed, p = 0.352). Finally, excavation lengths did not differ between cross-fostered and non-fostered animals (Figure 2G; ANCOVA, p = 0.075; Cohen's *d* = 0.53), and, if anything, the trend is in the opposite direction of expectation if a *P. maniculatus* parental environment influences the burrowing behavior of offspring. In summary, we found no differences in burrowing behavior after cross-fostering, consistent with there being a strong genetic component to the development of burrowing behavior.

Ontogeny of Burrow Construction Is *P. polionotus*-Dominant

We next tested the hypothesis that differences in the developmental onset of burrowing in juveniles share a common genetic basis with the well-characterized differences in adult burrow architecture [2–4] using a *P. polionotus* × *P. maniculatus* experimental cross (Figure 3A).

The development of burrowing behavior in first-generation (F_1) hybrids closely matches that in *P. polionotus* in each parameter examined, including the proportion of mice constructing burrows (Figure 3B; Fisher's exact test, p = 0.378; 10 of 11 mice [F_1] versus 16 of 22 mice [*P. polionotus*]), the proportion of mice constructing escape tunnels (Figure 3B; Fisher's exact test, p = 1.00; 4 of 11 mice [F_1] versus 8 of 22 mice [*P. polionotus*]), and the length of excavations (Figure 3C; ANCOVA, p = 0.115; Cohen's *d* = 0.68). Moreover, F_1 hybrids differ significantly from *P. maniculatus* in all of these measures of burrowing behavior: proportion of mice constructing burrows (Figure 3B; Fisher's exact test, p < 0.0001; 10 of 11 mice [F_1] versus 0 of 16 mice [*P. maniculatus*]), proportion of mice constructing escape tunnels (Figure 3B; Fisher's exact test, p = 0.019; 4 of 11 mice [F_1] versus 0 of 16 mice [*P. maniculatus*]), and length of excavations (Figure 3C; ANCOVA, p < 0.0001; Cohen's *d* = 2.50). This inheritance pattern indicates that the genetic underpinnings of precocious burrowing, a developmental trait, are *P. polionotus* dominant, consistent with the pattern of inheritance observed for adult burrowing behavior (F_1 hybrid adults build *P. polionotus*-like burrows with regard to both length and shape [2, 4]).

A *P. polionotus* Allele Affects Both Juvenile Onset and Adult Expression of Burrowing Behavior

To test whether developmental traits (namely, precocious burrow construction) and adult traits (long entrance tunnels and presence of an escape tunnel) are genetically linked, we generated 60 backcross (BC) hybrids. If traits have an independent genetic basis, they are expected to become uncoupled in this recombinant BC generation. We assessed burrowing performance for each BC hybrid at four time points: two juvenile (P21 and P24) and two adult (P61 and P64) trials (Figure 3D). We targeted the P21 and P24 time points because *P. polionotus* reached adult-like burrowing frequencies at this stage, but *P. maniculatus* did not (Figure 1B). Half of the BC hybrids (31 of 60) dug at least one juvenile burrow (at the P21 or P24 time point) and thus were scored as precocious burrowers, whereas the remaining half (29 of 60) completed no juvenile burrows and were scored as delayed burrowers. This segregation pattern is consistent with a single-locus effect, but the sample size is notably small.

To investigate the relationship between age at onset of burrowing and juvenile and adult excavation length, we ran a linear mixed-effect model with repeated-measures. Precociousness was a significant predictor of excavation length at both juvenile and adult stages (Figure 3E; p < 0.0001). We found that developmental onset of burrowing and adult excavation length co-segregated in recombinant BC hybrids (Figure 3F; p = 0.006), with precocious animals digging, on average, adult excavations that were 6.7 cm longer than those of delayed burrowers (Figure 3F). These data indicate that age at onset of burrowing (a developmental trait) and tunnel length variation (in adults) share a pleiotropic genetic basis, are influenced by linked genes, or both.

To test whether regions of the genome that are associated with adult burrowing behavior also influence onset of burrowing in juveniles, we genotyped BC mice at four unlinked single nucleotide polymorphisms (SNPs) previously associated with differences in adult burrow structure [4]. We then ran a repeated-measures linear mixed-effect model for each marker (Figure S2). We found that inheritance of a *P. polionotus* allele on linkage group 2 was significantly associated with variation in burrowing behavior at both juvenile and adult stages (Figure 4A; linear mixed-effect model, p = 0.02, post hoc Benjamini-Hochberg correction with 10% false discovery rate [FDR]). BC juveniles inheriting a single *P. polionotus* allele at this marker were 25.5% more likely to dig burrows precociously than those homozygous for the *P. maniculatus* allele (Figure 4B; mean precociousness score, MM = 0.392 ± 0.088 SEM, n = 30; PM = 0.647 ± 0.089 SEM, n = 29; Fisher's exact test, p = 0.044). Hybrids carrying a *P. polionotus* allele also dug longer excavations as juveniles (Figure 4C; MM = 8.73 cm ± 1.05 SEM, n = 30; PM = 12.75 cm ± 1.07 SEM, n = 29; linear mixed-effect model, p = 0.0107) and as adults (Figure 4D; MM = 22.15 cm ± 1.78 SEM, n = 30; PM = 28.05 cm ± 1.83 SEM, n = 28; linear

(E) Excavation length at juvenile and adult time points. Shading indicates whether each individual was a precocious burrower (i.e., at least one complete burrow dug at P21 or P24) or delayed burrower (i.e., no burrows dug at P21 or P24). Trait means for each group are shown at both time points, with error bars indicating ±1 SEM. Data were analyzed using a linear mixed-effect model with repeated measures.

(F) Average adult excavation length of BC hybrids that, as juvenile burrowers, were either delayed or precocious. Black lines indicate the means for each group. Significance levels are indicated as follows: ns (not significant), p ≥ 0.05; *p ≤ 0.05; **p ≤ 0.01; ***p ≤ 0.001.

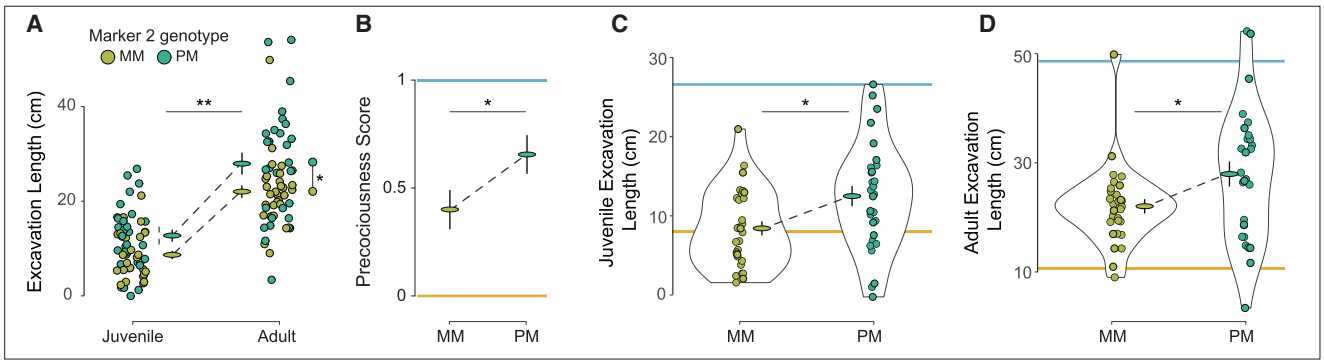


Figure 4. Effect of the Dominant *P. polionotus* Allele on Linkage Group 2 on Burrowing Behavior in Backcross Hybrids

Genotype at marker 2 is significantly associated with (A) variation in excavation length across life stages, (B) precociousness score across two juvenile trials (1, mouse dug at least one discrete burrow at P21 and/or P24 behavior trials; 0, mouse dug no burrow at P21 or P24), (C) average juvenile excavation length, and (D) average adult excavation length. Genotypes for 59 BC hybrids are either MM (homozygous for the *P. maniculatus allele*) or PM (heterozygous). For each genotype, trait means are plotted with error bars indicating ±1 SEM. The mean trait values for each parental species are plotted as horizontal bars (blue, *P. polionotus*; yellow, *P. maniculatus*). Parental species trait values are based on one trial per individual, aged P21–P24 (juveniles; *P. maniculatus*, n = 16; *P. polionotus*, n = 22) or >P60 (adults; *P. maniculatus*, n = 17; *P. polionotus*, n = 9). BC hybrid trait values are the average of two juvenile (A–C) or two adult (A and D) trials. Significance levels, determined by a linear mixed-effect model and Benjamini-Hochberg correction with 10% FDR (A), Fisher's exact test (B), or linear mixed-effect models (C and D) are indicated as follows: *$p \leq 0.05$; **$p \leq 0.01$. See also Figure S2 and Table S2.

mixed-effect model, p = 0.0303). Moreover, genotype at this marker explains variance in each behavior: precociousness score (6.4% phenotypic variance explained [PVE]), juvenile excavation length (10.9% PVE), and adult excavation length (8.2% PVE) [6]; the remaining unexplained variance in each trait could arise from environmental factors, additional genetic loci, or both. For each of the other markers examined, no significant relationships between genotype and phenotype were detected (Figure S2; p > 0.05), possibly due, in part, to the limited number of BC hybrids examined. Together, these data suggest that a gene, or linked genes, on linkage group 2 affects variation in burrowing behavior at different life stages.

DISCUSSION

> Huxley likes to speak of 'the three major problems of biology': that of causation, that of survival value and that of evolution—to which I should like to add a fourth, that of ontogeny. —Nikolaas Tinbergen (*On Aims and Methods of Ethology*)

Striking behavioral differences between closely related species can be a powerful resource for understanding the evolution of behavior and its mechanistic underpinnings—both major goals of biology. Behaviors are among the most complex phenotypes, and successfully teasing apart how species-specific differences evolve requires an integrative approach, as championed by Tinbergen [1]. More specifically, Tinbergen's 1963 landmark paper advocates for the addition of ontogeny to Huxley's existing framework for behavioral research [7].

Ontogeny, the study of how behavior changes across the life of an individual, can provide understanding that is not discernible using other approaches; for example, it can uncover unexpected ancestral state reconstructions and generate novel hypotheses (e.g., [8–11]) or expose underlying proximate mechanisms driving changes in behavior (e.g., [12–15]). In short, ontogeny informs and edifies each of Tinbergen's four questions and can provide novel insights into how behavior evolves.

Here, we focused on the ontogeny of burrow construction, an ecologically important behavior that varies dramatically between closely related species of North American *Peromyscus* rodents. Most species in this genus build small (<20 cm), simple burrows as adults, but one species, *P. polionotus*, has recently evolved a stereotyped burrowing behavior that results in a considerably longer burrow (>100 cm in the wild) consisting of an elongated entrance tunnel, a nest chamber, and a secondary tunnel that extends upward from the nest toward the soil surface. This second tunnel does not penetrate the soil surface except during emergency evacuation and thus is often referred to as an escape tunnel [2–4, 16–19] (Figure 1A). The burrows of *P. polionotus* have inspired studies of phylogenetic history [3], genetic mechanisms of behavior [2, 4], and speculations of adaptive function—namely, that *P. polionotus* burrows may provide refuge from the elevated rates of predation that occur in open, exposed habitats (e.g., [20, 21]). However, the ontogeny of the behavior—the last of Tinbergen's four questions—remained unexamined until now.

We report on how the final product of digging behavior—the extended phenotype [22], or burrow—originates and progresses during the post-natal development of two sister species of *Peromyscus* with dramatically different adult burrow architectures. We first find that *P. polionotus* are precocious with respect to burrow construction, building their first burrows 10 days earlier in development than *P. maniculatus*. This is surprising given that *P. maniculatus* is larger, tends to reach developmental milestones earlier, and outperforms age-matched *P. polionotus* in a wheel-running assay. These results suggest that *P. polionotus* has evolved a life history change—a precocious expression of behavior—that is most likely specific to burrow construction.

We also examined the shape of burrows produced by juvenile *Peromyscus* mice. We found that each species' characteristic

burrow architecture is intact in juveniles. This result suggests that in pure species, the neurobiological control of each component of the complete burrow architecture (frequency of burrow construction, entrance tunnel, and escape tunnel) is expressed together throughout life. This result is especially surprising in light of previous work showing that the genetic control of adult burrow construction in *P. polionotus* is modular [4]. Although the shape of juvenile burrows is similar to that of adult burrows, they are smaller in overall size, most likely due to the energetic cost of burrowing.

Using a cross-fostering experiment, we next tested whether these juvenile burrowing traits were primarily learned postnatally or were driven by interspecific genetic differences. It is important to note, however, that our experiments cannot rule out prenatal maternal effects (e.g., [23]). We found that cross-fostering results do not differ if single or multiple pups are transferred to heterospecific parents, suggesting that there is no measurable effect of sibling's genotype on juvenile behavior. We report that all aspects of species-specific burrowing behavior are preserved in cross-fostered individuals of both species, demonstrating that juvenile expression of burrowing behavior most likely has a strong genetic basis.

Finally, we examined the genetic underpinnings of behavioral ontogeny in hybrids of *P. polionotus* and *P. maniculatus* using a genetic cross. We found that a developmental trait (precocious onset of burrowing) and an adult trait (long tunnels characteristic of adult *P. polionotus* burrows) are genetically dominant and co-inherited, both at the level of phenotypic co-variation and with respect to a specific genetic marker. This is a surprising result, as behavior need not be correlated across life stages; indeed, many behaviors are expressed at only one stage. Although a well-powered genetic mapping study of burrowing development would be necessary to fully describe the genetic architecture of precocious burrowing, our data point to a shared—most likely pleiotropic—genetic influence on burrowing behavior that acts across juvenile and adult life stages.

These results have implications for the evolution of burrowing behavior. First, pleiotropy (or linkage of multiple causal mutations) can facilitate or inhibit evolution. On one hand, pleiotropy can produce effects that are not directly selected for (and potentially even harmful), but that are nevertheless secondarily "dragged along" by evolution [24, 25]. On the other hand, because changes in several traits are often involved during adaptation to a new environment [26–28], co-inheritance of groups of phenotypes (e.g., by pleiotropy or linkage) can expedite adaptation [29–31]. Indeed, a common experimental outcome is to map multiple traits to a shared genomic region [32–37], and this genetic architecture can affect how evolution proceeds.

Related, these findings make it difficult to identify the precise phenotypic targets of selection, if any. Although variation in adult burrows can affect fitness [21, 38], juvenile burrowing behavior may also be a target of selection. For example, natural selection for earlier burrowing in *P. polionotus* may reflect (1) its open habitat [16], which may expose young mice to predation and thus increase the survival value of burrowing, or (2) a form of "play" during a critical period of motor development [39–41]. Our results, which implicate a broadly acting pleiotropic genetic mechanism, highlight the challenge in identifying which specific trait or traits have been selected—in this case, precocious juvenile burrowing, long adult burrows, or both.

All animals integrate signals of their internal state with environmental cues to make behavioral choices that affect their survival and reproduction. These choices are made in an ecological context that often differs between species, which may—through a process of evolution by natural selection—produce heritable differences between species in the tuning of innate internal states and behavioral drives. We hypothesize that tuning of behavioral drives (over evolutionary time) provides a parsimonious explanation for the shared genetic control of developmental timing and expression of adult behavior in *Peromyscus* burrow construction (although other neural mechanisms are possible). More specifically, species-specific genetic differences may produce heritable internal states that persist in individuals across life stages, leading *P. polionotus* mice to engage in burrowing behavior earlier in life and also more frequently as adults than *P. maniculatus*, whose innate drives are tuned differently. Divergent neural tuning has often been linked to variation in neuromodulators or their receptors, rather than to variation in the underlying circuitry (e.g., [42–45]). Our results raise the possibility that neuromodulators (and behavioral drives) may be involved in the evolution of burrowing in *Peromyscus* rodents, consistent with the accumulating evidence that neuromodulatory systems are a frequent substrate for behavioral diversity and evolution [46].

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Animal husbandry
  - *Peromyscus maniculatus*
  - *Peromyscus polionotus*
  - Cross-fostered mice
  - *P. polionotus* x *P. maniculatus* F_1 Hybrids
  - Backcross Hybrids
- METHOD DETAILS
  - Burrowing behavior: Parental species and F_1 hybrids
  - Burrowing behavior: Backcross hybrids
  - Wheel-running behavioral trials
  - Genotyping
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Burrow Measurements
  - Statistical Analyses

SUPPLEMENTAL INFORMATION

Supplemental Information includes two figures and two tables and can be found with this article online at https://doi.org/10.1016/j.cub.2017.10.061.

AUTHOR CONTRIBUTIONS

H.C.M., N.L.B., and H.E.H. conceived and designed experiments. H.C.M., N.L.B., and Y.L.P. performed experiments. H.C.M. and N.L.B. analyzed the data. H.C.M. and H.E.H wrote the paper.

ACKNOWLEDGMENTS

We thank S. Scalia, J. Mason, and the Hoekstra Lab for assistance with behavioral assays and animal husbandry; A. Bendesky, B. de Bivort, C. Dulac, J. Losos, B. Ölveczky, H. Fisher, B. Peterson, M. Muñoz, and J. Weber for helpful discussions and comments on the manuscript; Harvard's Office of Animal Resources, particularly J. Rocca, for animal husbandry; and S. Worthington for statistical consultation. This research was funded by a Chapman Memorial Scholarship, a National Science Foundation (NSF) Graduate Research Fellowship, a Doctoral Dissertation Improvement Grant (NSF 1209753), and an American Fellowship from the American Association of University Women to H.C.M.; a Natural Sciences and Engineering Research Council of Canada Graduate Fellowship to N.L.B.; a Harvard College Research Program award and a Harvard Museum of Comparative Zoology Grant for Undergraduate Research to Y.L.P.; and a Beckman Young Investigator Award to H.E.H. H.E.H. is an Investigator of the Howard Hughes Medical Institute.

Received: June 14, 2017
Revised: August 28, 2017
Accepted: October 24, 2017
Published: November 30, 2017

REFERENCES

1. Tinbergen, N. (1963). On aims and methods of ethology. Z. Tierpsychol. *20*, 410–433.
2. Dawson, W.D., Lake, C.E., and Schumpert, S.S. (1988). Inheritance of burrow building in *Peromyscus*. Behav. Genet. *18*, 371–382.
3. Weber, J.N., and Hoekstra, H.E. (2009). The evolution of burrowing behaviour in deer mice (genus *Peromyscus*). Anim. Behav. *77*, 603–609.
4. Weber, J.N., Peterson, B.K., and Hoekstra, H.E. (2013). Discrete genetic modules are responsible for complex burrow evolution in *Peromyscus* mice. Nature *493*, 402–405.
5. Colonnese, M.T., Stallma, E.L., and Berridge, K.C. (1996). Ontogeny of action syntax in altricial and precocial rodents: grooming sequences of rat and guinea pig pups. Behaviour *133*, 1165–1195.
6. Broman, K.W., and Sen, S. (2009). A Guide to QTL Mapping with R/qtl (Springer).
7. Huxley, J. (1942). Evolution: The Modern Synthesis (John Wiley & Sons).
8. Dial, K.P. (2003). Wing-assisted incline running and the evolution of flight. Science *299*, 402–404.
9. Heers, A.M., and Dial, K.P. (2012). From extant to extinct: locomotor ontogeny and the evolution of avian flight. Trends Ecol. Evol. *27*, 296–305.
10. Moczek, A.P., Cruickshank, T.E., and Shelby, A. (2006). When ontogeny reveals what phylogeny hides: gain and loss of horns during development and evolution of horned beetles. Evolution *60*, 2329–2341.
11. Campbell, P., Pasch, B., Warren, A.L., and Phelps, S.M. (2014). Vocal ontogeny in neotropical singing mice (*Scotinomys*). PLoS ONE *9*, e113628.
12. Schulz, D.J., and Robinson, G.E. (2001). Octopamine influences division of labor in honey bee colonies. J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol. *187*, 53–61.
13. Gardner, T.J., Naef, F., and Nottebohm, F. (2005). Freedom and rules: the acquisition and reprogramming of a bird's learned song. Science *308*, 1046–1049.
14. Moon-Fanelli, A. (2011). The ontogeny of expression of communicative genes in coyote-beagle hybrids. Behav. Genet. *41*, 858–875.
15. Schneiderova, I. (2014). Vocal repertoire ontogeny of the captive Asian house shrew *Suncus murinus* suggests that the male courtship call develops from the caravanning call of the young. Acta Theriol. (Warsz.) *59*, 149–164.
16. Sumner, F.B., and Karol, J.J. (1929). Notes on the burrowing habits of *Peromyscus polionotus*. J. Mammal. *10*, 213–215.
17. Hayne, D.W. (1936). Burrowing habits of *Peromyscus polionotus*. J. Mammal. *17*, 420–421.
18. Rand, A.L., and Host, P. (1942). Mammal notes from Highland County, Florida. Bull. Am. Mus. Nat. Hist. *80*, 1–21.
19. Ivey, R.D. (1949). Life history notes on three mice from the Florida east coast. J. Mammal. *30*, 157–162.
20. Blair, W.F. (1951). Population structure, social behavior, and environmental relations in a natural population of the beach mouse (*Peromyscus polionotus leucocephalus*). Contrib. Lab. Vertebr. Biol. Univ. Mich. *48*, 1–47.
21. Jackson, T.P. (2000). Adaptation to living in an open arid environment: lessons from the burrow structure of the two southern African whistling rats, *Parotomys brantsii* and *P. littledalei*. J. Arid Environ. *46*, 345–355.
22. Dawkins, R. (1982). The Extended Phenotype: The Gene as the Unit of Selection (Freeman).
23. Francis, D.D., Szegda, K., Campbell, G., Martin, W.D., and Insel, T.R. (2003). Epigenetic sources of behavioral differences in mice. Nat. Neurosci. *6*, 445–446.
24. Cooper, T.F., Ostrowski, E.A., and Travisano, M. (2007). A negative relationship between mutation pleiotropy and fitness effect in yeast. Evolution *61*, 1495–1499.
25. Stern, D.L. (2013). The genetic causes of convergent evolution. Nat. Rev. Genet. *14*, 751–764.
26. Fisher, R.A. (1930). The Genetical Theory of Natural Selection (Oxford University Press).
27. Orr, H.A. (2000). Adaptation and the cost of complexity. Evolution *54*, 13–20.
28. Schluter, D.S. (2000). The Ecology of Adaptive Radiation (Oxford University Press).
29. Kirkpatrick, M., and Barton, N. (2006). Chromosome inversions, local adaptation and speciation. Genetics *173*, 419–434.
30. Hoffmann, A.A., and Rieseberg, L.H. (2008). Revisiting the impact of inversions in evolution: from population genetic markers to drivers of adaptive shifts and speciation? Annu. Rev. Ecol. Evol. Syst. *39*, 21–42.
31. Albert, A.Y.K., Sawaya, S., Vines, T.H., Knecht, A.K., Miller, C.T., Summers, B.R., Balabhadra, S., Kingsley, D.M., and Schluter, D. (2008). The genetics of adaptive shape shift in stickleback: pleiotropy and effect size. Evolution *62*, 76–85.
32. Mills, M.G., Greenwood, A.K., and Peichel, C.L. (2014). Pleiotropic effects of a single gene on skeletal development and sensory system patterning in sticklebacks. Evodevo *5*, 5–10.
33. Hawthorne, D.J., and Via, S. (2001). Genetic linkage of ecological specialization and reproductive isolation in pea aphids. Nature *412*, 904–907.
34. Hall, M.C., Basten, C.J., and Willis, J.H. (2006). Pleiotropic quantitative trait loci contribute to population divergence in traits associated with life-history variation in *Mimulus guttatus*. Genetics *172*, 1829–1844.
35. Protas, M., Tabansky, I., Conrad, M., Gross, J.B., Vidal, O., Tabin, C.J., and Borowsky, R. (2008). Multi-trait evolution in a cave fish, *Astyanax mexicanus*. Evol. Dev. *10*, 196–209.
36. Joron, M., Frezal, L., Jones, R.T., Chamberlain, N.L., Lee, S.F., Haag, C.R., Whibley, A., Becuwe, M., Baxter, S.W., Ferguson, L., et al. (2011). Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. Nature *477*, 203–206.
37. Miller, C.T., Glazer, A.M., Summers, B.R., Blackman, B.K., Norman, A.R., Shapiro, M.D., Cole, B.L., Peichel, C.L., Schluter, D., and Kingsley, D.M. (2014). Modular skeletal evolution in sticklebacks is controlled by additive and clustered quantitative trait loci. Genetics *197*, 405–420.
38. Hansell, M.H. (2005). Animal Architecture (Oxford University Press).
39. Fagen, R. (1981). Animal Play Behavior (Oxford University Press).
40. Burghardt, G.M. (2005). The Genesis of Animal Play (MIT Press).
41. Byers, J.A., and Walker, C. (1995). Refining the motor training hypothesis for the evolution of play. Am. Nat. *146*, 25–40.
42. Insel, T.R., Gelhard, R., and Shapiro, L.E. (1991). The comparative distribution of forebrain receptors for neurohypophyseal peptides in monogamous and polygamous mice. Neuroscience *43*, 623–630.

43. Insel, T.R., and Shapiro, L.E. (1992). Oxytocin receptor distribution reflects social organization in monogamous and polygamous voles. Proc. Natl. Acad. Sci. USA *89*, 5981–5985.

44. Marder, E., Calabrese, R.L., Nusbaum, M.P., and Trimmer, B. (1987). Distribution and partial characterization of FMRFamide-like peptides in the stomatogastric nervous systems of the rock crab, *Cancer borealis*, and the spiny lobster, *Panulirus interruptus*. J. Comp. Neurol. *259*, 150–163.

45. Verley, D.R., Doan, V., Trieu, Q., Messinger, D.I., and Birmingham, J.T. (2008). Characteristic differences in modulation of stomatogastric musculature by a neuropeptide in three species of *Cancer* crabs. J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol. *194*, 879–886.

46. Bendesky, A., and Bargmann, C.I. (2011). Genetic contributions to behavioural diversity at the gene-environment interface. Nat. Rev. Genet. *12*, 809–820.

47. Dawson, W.D. (1964). Fertility and size inheritance in a *Peromyscus* species cross. Evolution *9*, 44–55.

48. Vrana, P.B., Fossella, J.A., Matteson, P., del Rio, T., O'Neill, M.J., and Tilghman, S.M. (2000). Genetic and epigenetic incompatibilities underlie hybrid dysgenesis in *Peromyscus*. Nat. Genet. *25*, 120–124.

49. Dewsbury, D.A. (1980). Wheel-running behavior in 12 species of muroid rodents. Behav. Processes *5*, 271–280.

50. Venables, W.N., and Ripley, B.D. (2002). Modern Applied Statistics with S, Fourth Edition (Springer).

51. Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. J. Stat. Softw. *67*, 1–48.

52. Singmann, H., Bolker, B., Westfall, J., and Aust, F. (2016). afex: analysis of factorial experiments. R package version 0.16-1. https://github.com/singmann/afex.

53. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. B *57*, 289–300.

54. Kuznetsova, A., Brockhoff, P.B., and Christensen, R.H.B. (2016). lmerTest: tests in linear mixed effects models. R package version 2.0-33.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
Restriction enzyme *HinfI*	New England Biolabs	Cat#R0155S
Restriction enzyme *MseI*	New England Biolabs	Cat#R0525S
Experimental Models: Organisms/Strains		
Peromyscus polionotus subgriseus; PO stock	Peromyscus Genetic Stock Center	N/A
Peromyscus maniculatus bairdii; BW stock	Peromyscus Genetic Stock Center	RRID: MMR-RC_041477-MU
Oligonucleotides		
See Table S2	Integrated DNA Technologies	N/A
Software and Algorithms		
R/qtl	http://www.rqtl.org/	N/A
Other		
Polyurethane filling foam	HILTI	Cat#CF116
Small animal exercise wheel	Ware Manufacturing	Cat#03281
Wireless cyclocomputer	Cateye	Cat#CC-COM10W

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Hopi Hoekstra (hoekstra@oeb.harvard.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animal husbandry

We conducted experiments using captive *Peromyscus* strains mice kept under controlled laboratory conditions. All mice were housed in ventilated cages at 22° C on a 16:8 h light:dark cycle and provided food and water *ad libitum*. Breeding pairs and their litters were fed irradiated PicoLab Mouse Diet 20 5058 (LabDiet, St. Louis, MO) and virgin mice were fed irradiated LabDiet Prolab Isopro RMH 3000 5P75 after weaning. Animals were provided with cotton nesting material, corn cob bedding, and 3-sided red polycarbonate shelters. Juveniles were weaned at P24 into cages with at most four other animals (of the same sex and strain, unless otherwise noted). For all experiments, we used only offspring of experienced parents ($\geq$1 previous litter weaned). All procedures were approved by the Harvard University Institutional Animal Care and Use Committee (Protocol ID 27-09-1).

Peromyscus maniculatus

Peromyscus maniculatus bairdii (BW stock) were originally acquired from the *Peromyscus* Genetic Stock Center (University of South Carolina, Columbia SC, USA); this outbred line was derived from wild-caught ancestors in 1948 and has been laboratory-housed since capture. We formed eight breeding pairs using unrelated adults and checked daily for the presence of new pups. We tested 39 juveniles (20 females, 19 males) and 17 adults (8 females, 9 males). See Figure S1 for measurements of body mass for this species across development.

Peromyscus polionotus

Peromyscus polionotus subgriseus (PO stock) were acquired from the *Peromyscus* Genetic Stock Center; this outbred line was derived from wild-caught ancestors in 1952. We formed nine breeding pairs using unrelated adults and checked daily for the presence of new pups. We tested 58 juveniles (24 females, 34 males) and 9 adults (3 females, 6 males). See Figure S1 for measurements of body mass for this species across development.

Cross-fostered mice

Age-matched ($\leq$48 hr age difference) *P. maniculatus* (n = 18; 8 females, 10 males) and *P. polionotus* (n = 16; 5 females, 11 males) pups were traded between experienced ($\geq$1 previous litter) heterospecific breeding pairs 24-48 hr after birth. To test for effects of parents versus siblings on the behavior of the test animal(s), we used two fostering paradigms: pups were fostered as either individuals (one pup traded between litters, such that the fostered pup had heterospecific siblings and heterospecific parents) or as litters

(entire litters traded between breeding pairs, such that pups had heterospecific foster parents but conspecific siblings). Because burrowing performance of both singly and group cross-fostered animals did not differ (ANCOVA; *P. polionotus*: age p = 0.010, foster treatment p = 0.880, age x treatment interaction p = 0.677; *P. maniculatus:* age p = 0.006, foster treatment p = 0.807, treatment x age interaction p = 0.853), we grouped these data together for subsequent analyses comparing fostered and non-fostered animals. Following weaning, juveniles were housed with mixed-sex siblings (biological or foster) until completion of behavioral trials. We measured the burrowing behavior of each resultant juvenile at a single time point (during P19-P31).

P. polionotus x *P. maniculatus* F_1 Hybrids

We produced F_1 hybrids by crossing *P. maniculatus* dams to *P. polionotus* sires. Due to genomic imprinting in these species, our cross design for production of F_1 hybrids was limited to one direction [47, 48]. Thus, this cross design excludes any *P. polionotus* maternal effects acting in favor of *P. polionotus*-like burrowing behavior. We formed two breeding pairs using unrelated adults and checked daily for the presence of new pups. Eleven F_1 hybrids were tested (5 females, 6 males). Weanlings were subsequently housed with their mixed-sex littermates until completion of experiments.

Backcross Hybrids

We generated 60 backcross hybrids (31 females, 29 males) by crossing F_1 hybrids to *P. maniculatus* mates (Figure 3A). Both male and female F_1 hybrids were backcrossed to *P. maniculatus* (reciprocal pairings); 22 animals were produced from an F_1 dam and 38 from an F_1 sire. We weaned animals after their last juvenile burrow test (P24), and weanlings were subsequently housed with four other age-matched, same-sex BC mice. Juvenile mice were weighed prior to testing at both P21 and P24 time points (female average weight: 10.08 g ± 0.18 SEM; male average weight: 9.99 g ± 0.27). These same BC mice were weighed again prior to adult testing at both P61 and P64 time points (female average weight: 15.20 g ± 0.28 SEM; male average weight: 16.57 g ± 0.47).

METHOD DETAILS

Burrowing behavior: Parental species and F_1 hybrids

We tested burrowing behavior in a total of 142 juvenile and 26 adult mice in large, indoor sand-filled arenas as previously described [3, 4], except duration was reduced from 48 hr to 14-17 hr (i.e., one 8-hour dark cycle followed by 6-9 hr of light) for juveniles. Briefly, we released animals into 1.2 × 1.5 × 1.1 m enclosures filled with approximately 700 kg of moistened, hard-packed premium playground sand (Pharma-Serv), under otherwise normal housing conditions. We tested juveniles once, singly, without previous experience, and thus our experiment targeted innate behavior and not learned ability.

We tested mice of both species at postnatal ages P19, P21, P23, P24, P25, P27, P29 and P31. Because of the species' early onset of burrowing behavior, we tested additional *P. polionotus* individuals at P17, the earliest possible age to separate a juvenile from its mother. We tested F_1 hybrids at P19, P21, and P24. Thus, we constructed a developmental time series for each species during key stages of motor and behavioral development.

Burrowing behavior: Backcross hybrids

We characterized both juvenile and adult burrowing behavior of 60 backcross mice, collecting developmental and adult phenotypes in the same individuals: each BC animal was tested four times in total, at juvenile ages P21 and P24, and adult ages P61 and P64 (apart from one individual that died prior to adult testing). Enclosure area was reduced by half (i.e., to 0.6 × 1.5 × 1.1 m) for assaying both juvenile and adult backcross individuals to accommodate the large number of animals being tested.

Wheel-running behavioral trials

To compare the ontogeny of a second motor behavior (and general activity level) between species, we performed a standardized wheel running assay [49]. We tested naive, juvenile *P. maniculatus* (n = 43; 15 females, 28 males) and *P. polionotus* (n = 40; 13 females, 27 males) at P17-P31. We also tested 10 adults (5 females, 5 males) of each species. After 4 hr of home cage habituation to the wheel (Ware Manufacturing, Phoenix, AZ), we recorded 90 min of wheel running activity with a CC-COM10W wireless bike computer (Cateye, Osaka, Japan). *Peromyscus* show strongly nocturnal patterns of wheel running [49], thus we performed all tests during the first 4 hr of the dark cycle. All animals were weighed prior to testing (juvenile *P. polionotus*: 8.98 g ± 0.29 SEM; juvenile *P. maniculatus*: 10.53 g ± 0.31; adult *P. polionotus*: 14.60 g ± 0.45; adult *P. maniculatus*: 16.46 g ± 0.46).

Genotyping

We genotyped the BC population (n = 60) at four markers corresponding to known loci underlying adult burrowing behavior (identified in [4]) using species-specific restriction fragment length polymorphism (RFLP) differences. We designed all four assays such that the *P. polionotus* allele contained a restriction enzyme cut site, whereas the *P. maniculatus* allele did not. We performed PCR with a Taq DNA Polymerase kit (QIAGEN) and custom primers (Integrated DNA Technologies; Table S2). We verified that the selected RFLPs were fixed between species by Sanger sequencing of PCR amplicons of four unrelated individuals of each species, as well as the *P. maniculatus* and F_1 parents of the backcross (BigDye Terminator v3.1 Cycle Sequencing Kit, Life Technologies). PCR products were digested with restriction enzymes (New England Biolabs, Ipswich MA; Table S2), separated by gel electrophoresis (with Quick-Load 100bp DNA Ladder (New England Biolabs, Ipswich MA) as a size reference), and genotypes were called based on

the resultant banding pattern. All BC animals inherit at least one *P. maniculatus* allele; therefore, we interpreted the presence of a second smaller fragment (of appropriate size) as evidence of a *P. polionotus* allele.

QUANTIFICATION AND STATISTICAL ANALYSIS

Burrow Measurements

To quantify burrow construction, at the conclusion of each trial, we inspected enclosures for any excavations, which were qualitatively characterized as either burrows (comprised of $\geq$ 1 tunnel plus a nest area) or divots (broad cup-shaped vertical diggings < 10 cm; see Figure 1F). Next, we injected unoccupied burrows with polyurethane insulation foam (Hilti, Schaan, Liechtenstein) as previously described [3, 4]. Dried polyurethane casts were numerically coded, and the lengths of burrow components (entrance tunnel, nest chamber, and escape tunnel if present) were later measured from the casts by a researcher blind to animal identity. Lengths of divots were measured directly in the enclosures.

Statistical Analyses

Behavior Development

To disentangle effects of age and species on burrowing behavior, we employed several statistical tests. We first tested for effects of age and species on burrowing behavior as well as for effects of sex, postnatal litter size, enclosure, and foster status at the intraspecific level using ANCOVA. When significant effects were found, we used Tukey's HSD to compare means. To indicate effect sizes, we calculated Cohen's *d* directly from the data (t tests) or from residuals after regressing out age (ANCOVAs). Because we did not detect statistical differences between treatments, singly cross-fostered individuals and litter-fostered animals were pooled (fostering details above). We used Fisher's exact test to evaluate the frequencies of burrow and escape tunnel digging between different groups. Two *P. polionotus* individuals that appeared in poor health (age > 23 days) were excluded from all analyses. All statistical analyses were performed in R (version 3.2.3).

Genetic Cross

To investigate the relationship between precociousness in juveniles and excavation length in adults (i.e., phenotypic correlations in recombinant BC hybrids), we first ran a linear mixed-effect model with repeated-measures. To identify which random variables to include in the mixed models, we used *stepAIC* {MASS} [50] to select the best-fit model by AIC comparison. We included the following variables as possible covariates: sex, body mass, trial number, enclosure number, cross direction, dam ID, sire ID, as well as maternal and paternal grandparent ID (i.e., family structure). The best-fit model for excavation length included sex, body mass, and maternal grandparent ID. We therefore constructed a repeated-measures linear mixed-effect model of total excavation length using age, precociousness, and age:precociousness as fixed effects and mouse ID, sex, body mass, and maternal grandparent ID as random effects. Post hoc, we tested for effects of precocious burrowing on adult excavation length only, with the best-fit model including body mass and trial number. We therefore constructed a repeated-measures linear mixed-effect model of adult excavation length with precociality as a fixed explanatory variable and mouse ID, body mass, and trial number as random variables. All mixed models were run using *lmer* {lme4} [51], and p values for all fixed effects were calculated using *mixed* {afex} [52].

Marker associations

To evaluate the relationship between genotype and burrowing phenotype in BC hybrids, we ran a repeated-measures linear mixed-effect model. We used stepwise AIC model comparison with *stepAIC* {MASS} [50] to determine which covariates to include. The best-fit models for markers 1-3 (examining total excavation length) included age, genotype, and age:genotype as fixed effects and mouse ID, sex, body mass, and maternal grandparent ID as random variables. The best-fit model for marker 4 (examining escape tunnel length) included age, genotype, and age:genotype as fixed effects and mouse ID and body mass as random variables. All mixed models were run with *lmer* {lme4} [51], and p values for fixed effects were calculated using *mixed* {afex} [52]. To further evaluate the significance of these genotype-phenotype associations, we implemented Benjamini-Hochberg corrections [53] for multiple comparisons with 10% FDR. Post hoc tests for associations between marker 2 (*Chrm5*) and specific phenotypes included additional linear mixed-effect models (juvenile excavation length; adult excavation length) and a Fisher's exact test (precociousness score). We estimated p values for post hoc models using *summary* {lmerTest} [54]. Effect sizes of genotype on each phenotype are reported as genotype-specific phenotype averages, plotted with standard error of the mean. Calculations of percent variance explained for each trait are based on marker regression mapping [6].