LETTER

doi:10.1038/nature11816

Discrete genetic modules are responsible for complex burrow evolution in *Peromyscus* mice

Jesse N. Weber[1]†, Brant K. Peterson[1,2] & Hopi E. Hoekstra[1,2]

Relative to morphological traits, we know little about how genetics influence the evolution of complex behavioural differences in nature[1]. It is unclear how the environment influences natural variation in heritable behaviour[2], and whether complex behavioural differences evolve through few genetic changes, each affecting many aspects of behaviour, or through the accumulation of several genetic changes that, when combined, give rise to behavioural complexity[3]. Here we show that in nature, oldfield mice (*Peromyscus polionotus*) build complex burrows with long entrance and escape tunnels, and that burrow length is consistent across populations, although burrow depth varies with soil composition. This burrow architecture is in contrast with the small, simple burrows of its sister species, deer mice (*P. maniculatus*). When investigated under laboratory conditions, both species recapitulate their natural burrowing behaviour. Genetic crosses between the two species reveal that the derived burrows of oldfield mice are dominant and evolved through the addition of multiple genetic changes. In burrows built by first-generation backcross mice, entrance-tunnel length and the presence of an escape tunnel can be uncoupled, suggesting that these traits are modular. Quantitative trait locus analysis also indicates that tunnel length segregates as a complex trait, affected by at least three independent genetic regions, whereas the presence of an escape tunnel is associated with only a single locus. Together, these results suggest that complex behaviours—in this case, a classic 'extended phenotype'[4]—can evolve through multiple genetic changes each affecting distinct behaviour modules.

Animal architectures, such as beehives, bird nests, spider webs, termite mounds and rodent burrows, are remarkably diverse traits that can evolve through natural selection. Despite their great diversity, these extended phenotypes have similarities: they seem to be constructed through largely unlearned motor patterns; they are often consistent within a species (or population); and, when architectures differ, these differences reflect important fitness-related functions in the wild[5]. Thus, genetic changes are predicted to contribute to the evolution of different architectures, even between closely related species; however, biologists have long questioned how genetic changes can lead to the evolution of distinct behaviours[6]. Because animal architectures can be precisely measured—similarly to morphological traits, which have been successfully dissected genetically[7,8]—these are excellent traits for the genetic analysis of behaviour.

Many species of *Peromyscus* mice build burrows, which were initially described by natural historians working in the field[9–14]. However, these burrowing behaviours can also be studied in the laboratory[15,16], and our previous work showed that burrow differences among *Peromyscus* mice have a strong genetic component[17]. In particular, *P. polionotus* and *P. maniculatus* are recently diverged and interfertile sister species that build distinct burrows. *P. polionotus* is an open-field specialist and is restricted to the southeastern United States (Fig. 1a and Supplementary Fig. 1a), whereas *P. maniculatus* is a generalist, which inhabits prairie and forest habitats across much of North America (Supplementary Fig. 1c). Whereas most *Peromyscus* species build small, single-tunnel burrows (Supplementary Fig. 1d) or no burrows at all, *P. polionotus* construct unique burrows characterized by a long entrance tunnel that leads into a nest cavity, and a secondary tunnel that emanates from the nest and terminates just below the soil surface (Fig. 1b and Supplementary Fig. 1b)[9–11]. This secondary tunnel may serve several functions[11], most notably its use as an escape tunnel when predators (for example, snakes) invade the entrance tunnel (see Supplementary Video). Overall, the complex burrow architecture in *P. polionotus* is derived[17], and probably associated with adaptation to living in an exposed, open habitat[15].

To examine natural variation in burrow complexity, we measured burrow dimensions and soil composition (that is, percentage silt, sand and clay) across the range of *P. polionotus* (Fig. 1a and Supplementary

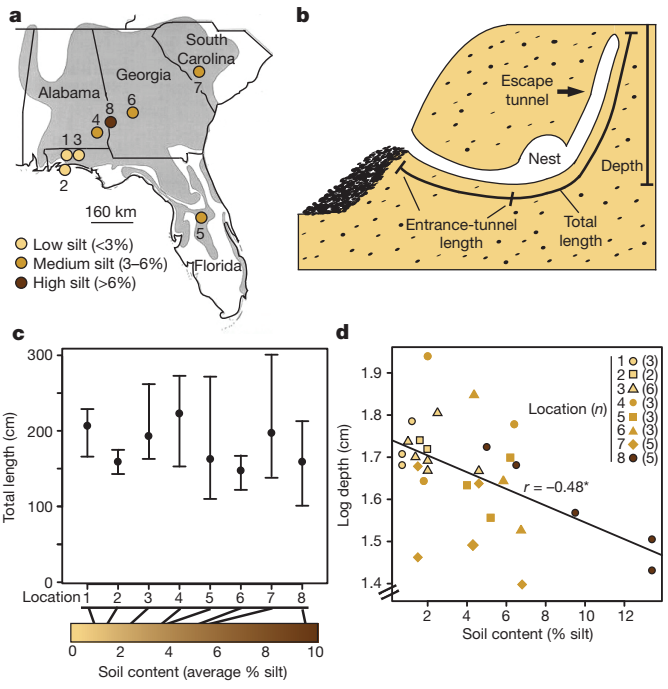

Figure 1 | Natural variation in *P. polionotus* burrows. **a**, Sampling of burrows at eight sites in the southeastern United States from across the range of *P. polionotus* (grey area). Average percentage of soil silt at each sampling site is provided. **b**, Diagram of a typical *P. polionotus* burrow showing the measures for entrance-tunnel length, total length and burrow depth, as well as a typical escape tunnel. **c**, Variation in total burrow length among sites (mean ± range; from **a**), which are ordered by increasing percentage of silt (left to right). **d**, Correlation between silt composition of soil and burrow depth (asterisk indicates Spearman correlation, rho = −0.48, $S = 6632$, $P = 0.01$). Each point represents a burrow, and shapes represent the eight different sampling sites (from **a**). The number of burrows measured at each site is shown in parentheses.

[1]Department of Organismic & Evolutionary Biology, Museum of Comparative Zoology, 26 Oxford Street, Cambridge, Massachusetts 02138, USA. [2]Department of Molecular & Cellular Biology, Center for Brain Science, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA. †Present address: Section of Integrative Biology, One University Station, University of Texas Austin, Texas 78712, USA.

Table 1; see Methods for details). Burrows generally had two long tunnels—an entrance and an escape tunnel (Fig. 1b)—and a mean total length of 181 cm (standard deviation (s.d.) of ± 53). Of the three soil variables measured, the percentage of silt was the most likely to influence soil compaction and thereby influence burrow dimensions[18]. Despite large differences in the percentage of silt among locations (Fig. 1c; Kruskal–Wallis (K–W) test, chi-squared = 19.42, degrees of freedom (df) = 7, $P = 0.01$), burrow length did not differ significantly among locations (K–W test, chi-squared = 6.59, df = 7, $P = 0.47$), nor did burrow length correlate with the percentage of silt measured at each burrow (Spearman test, rho = −0.16, $S = 5{,}226$, $P = 0.39$). In contrast, there was a significant, negative correlation between the percentage of silt in the soil at each burrow and burrow depth (Fig. 1d; Spearman test, rho = −0.48, $S = 6{,}632$, $P = 0.01$), suggesting that the burrows are shallower, but not shorter, when constructed in compact, silty soil[11] (Fig. 1d). Overall, natural burrow shape and length are remarkably conserved in wild *P. polionotus*, yet variation in soil composition affects burrow depth. We therefore focused our genetic analyses on burrow length.

To measure the genetic component of complex burrowing, we first assayed the burrowing behaviour of *P. polionotus* and *P. maniculatus* in a single soil type under controlled laboratory conditions (see Methods). Briefly, each assay involved placing a mouse in a large, sand-filled enclosure for 46 h (two full-night activity periods). Then, we removed the mouse and made a polyurethane cast of the burrow (Supplementary Fig. 1c, d), which we then measured (Fig. 1b). We tested each animal in three consecutive assays. Captive-reared mice, which had never been exposed to sand or allowed to burrow, recapitulated their natural burrowing behaviour in our assays. We found no significant effects of assay number, sex, age, mass or enclosure on total burrow length (data not shown). Statistical comparisons of *P. polionotus* and *P. maniculatus* burrows revealed significant differences in total burrow length (Welch's two-tailed *t*-test, $t = 3.24$, df = 20.98, $P < 0.01$), consistent with previous results[15,16]. Moreover, total burrow length is composed of two main parts: entrance-tunnel length, which differed significantly between species (Fig. 2b; Welch's two-tailed *t*-test, $t = 6.72$, df = 24.39, $P < 5 \times 10^{-3}$), and escape-tunnel length—only *P. polionotus* constructed escape tunnels (Fig. 2c). Thus, despite having no previous experience with either sand substrate or our enclosures, these species built consistently distinct burrows in the laboratory.

To determine the inheritance patterns of this behavioural variation, we characterized the genetic architecture of burrowing differences observed between *P. polionotus* and *P. maniculatus*. First, we crossed the two species and assayed the burrows of their hybrid offspring (Supplementary Methods and Fig. 2a). First-generation (F_1) hybrids built entrance tunnels that were significantly longer (after Bonferroni correction, $\alpha = 1.67 \times 10^{-2}$) than *P. maniculatus* (Welch's two-tailed *t*-test, $t = 5.34$, df = 20.7, $P = 2 \times 10^{-4}$). But F_1 entrance lengths did not differ from the *P. polionotus* parent (one sample *t*-test, $\mu = 14$ cm, $t = -0.14$, df = 12, $P = 0.89$) or from the *P. polionotus* population mean (Fig. 2b; Welch's two-tailed *t*-test, $t = 2.35$, df = 22.7, $P = 0.03$), and all F_1 animals constructed escape tunnels (Fig. 2c). Thus, the alleles contributing to burrow size and shape seem to segregate in a dominant fashion.

We next crossed F_1 animals with *P. maniculatus* to create a recombinant backcross (BC) generation ($n = 272$). BC mice constructed entrance tunnels that varied continuously in length between the parental extremes, but approximately one of eight ($n = 36$) of the BC mice built *P. polionotus*-length tunnels (Fig. 2b; >14 cm in length), suggesting that only a few loci are necessary to generate this behaviour. In contrast, half (46%) of the BC mice built escape tunnels (Fig. 2c). This inheritance pattern is consistent with the action of either a single major-effect locus or of multiple loci that interact to create a threshold effect[19], such that only some loci need to be co-inherited to cause the expression of a trait. Finally, tunnel lengths and the presence of escape tunnels (that is, tunnel number) are only weakly correlated in BC mice

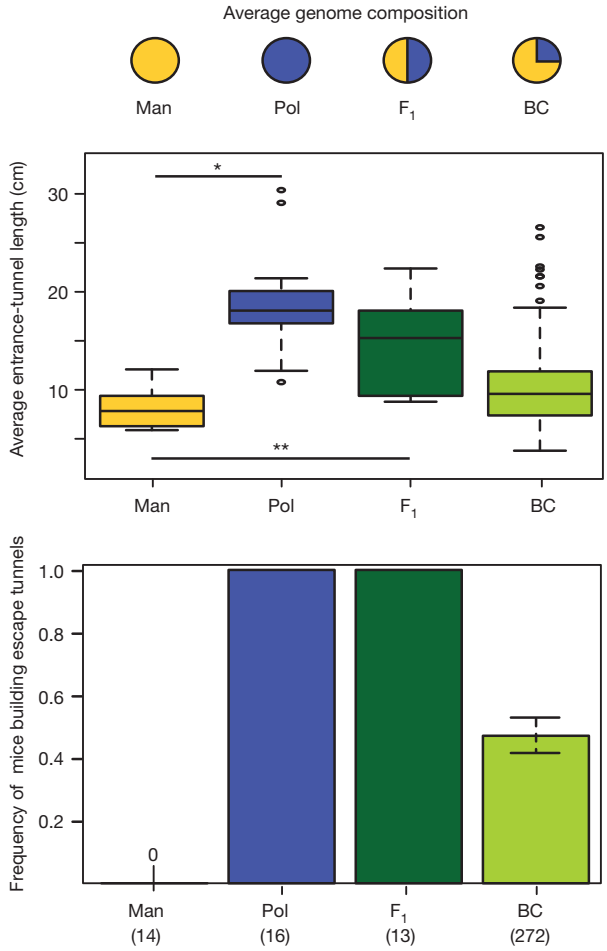

Figure 2 | Burrow variation across generations. **a**, Burrow dimensions of *P. maniculatus* (Man; yellow), *P. polionotus* (Pol; blue), F_1 hybrids (dark green) and progeny resulting from $F_1 \times$ *P. maniculatus* backcross (BC; light green). Pie charts depict average genome composition in each generation. Distributions of entrance-tunnel length (average of three trials for each individual tested) in the parental species, F_1 hybrids and BC animals are shown. Boxes represent interquartile ranges (median ± 2 s.d.). Significant *t*-tests, $*P = 5 \times 10^{-3}$, $** P = 2 \times 10^{-4}$. **b**, The frequency of escape-tunnel construction is shown for the same individuals. Error bars represent mean ± standard error of the mean (s.e.m.) Sample sizes are listed in parentheses below.

(Supplementary Table 2). Therefore, the complex burrows of *P. polionotus* comprise at least two separate behavioural modules, one for tunnel length and one for the presence of an escape tunnel.

To identify the chromosomal locations involved and quantify the phenotypic effects of loci that influence these behavioural modules, we used a quantitative trait locus (QTL) mapping approach. We first identified 526 diagnostic single nucleotide polymorphisms (SNPs) using a double-digest restriction-site-associated-DNA (ddRAD) method[20], and genotyped the 272 BC mice for which we measured burrowing behaviour (that is, entrance-tunnel length and the presence of an escape tunnel; see Methods for details). Consistent with simple inheritance patterns predicted by burrow phenotypes in hybrid animals, we identified three genomic regions that contributed to variation in entrance-tunnel length, and a single region associated with escape-tunnel construction (Fig. 3a and Supplementary Fig. 5). All four QTLs are unlinked and segregate on separate chromosomes. The entrance-tunnel-length QTLs have similar effect sizes, interact additively (as determined by both a lack of epistasis between loci and graphical comparisons; Fig. 3b and Supplementary Methods), and together explain ~15% of tunnel-length variation (Table 1). Because approximately 24% of the variation in burrow length is likely to have a

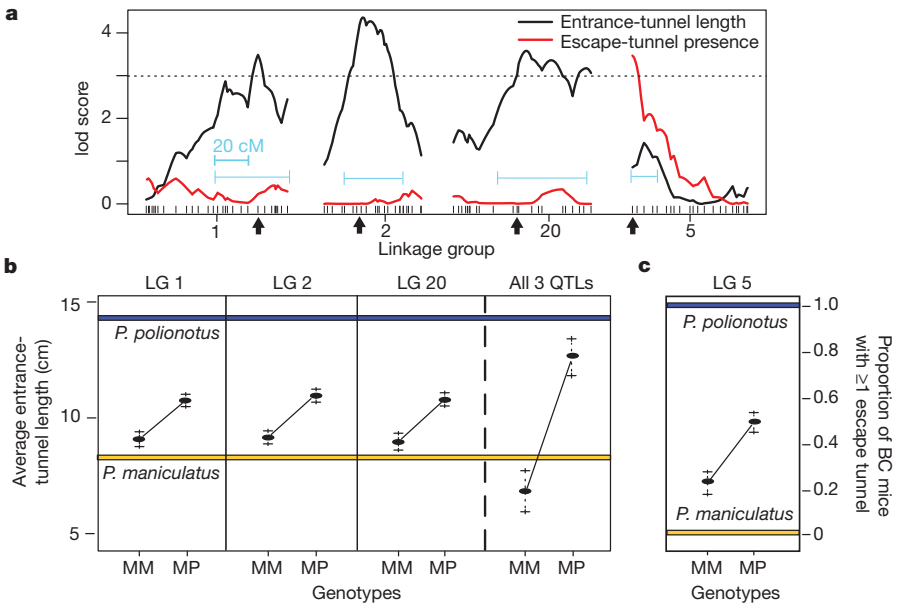

Figure 3 | QTL analysis of burrow variation. **a**, Linkage groups (LGs) 1, 2 and 20 harbour QTLs associated with log-transformed average entrance-tunnel length (black line). Linkage group 5 contains a single QTL associated with escape-tunnel presence (red line). Dotted line represents log odds ratio (lod) significance threshold (genome-wide $\alpha = 0.05$, lod = ~3.0). 1.5-lod confidence intervals and scale in centimorgans (cM) are shown in light blue. Dashes indicated genetic markers, and black arrows indicate markers used to define each QTL peak (used in **b**). **b**, Phenotypic effect of individual and combined QTLs (linkage groups 1, 2 and 20) on entrance-tunnel length in 272 BC mice. **c**, Proportion of BC animals that construct escape tunnels for each of the two genotypes. All error bars represent mean ± s.e.m. Blue and yellow lines represent average phenotype of the parents (pure species) used to found the cross. Genotypes are either homozygous *P. maniculatus* (MM) or heterozygous *P. maniculatus/polionotus* (MP).

genetic basis (as measured by the repeatability of burrow length in *P. polionotus*)[17], the QTLs that we report may explain more than half of the genetic variation for this trait (but see ref. 21). Even more notable, each *P. polionotus* allele increases entrance-tunnel length by, on average, 3 cm (Fig. 3b). Furthermore, BC individuals with all three *P. maniculatus* alleles versus those with one *P. polionotus* allele at each locus (that is, heterozygotes) nearly recapitulate the difference in entrance-tunnel length between the two species (that is, the mean phenotype of the parents used in the cross). Although it is possible that each QTL may harbour multiple, possibly interacting, mutations, these results are consistent with only a few genes (as few as three) being needed to explain the continuous tunnel-length variation observed in BC mice.

By contrast, QTL analysis identified only a single locus for the presence of escape tunnels on linkage group 5. Although this QTL explains only 6% of the escape-tunnel variance, BC individuals that inherited one *P. polionotus* allele at this locus were 30% more likely on average to build an escape tunnel than individuals with only *P. maniculatus* alleles (Fig. 3c). Given that the QTL is located at the end of a linkage group, it is possible that the causative mutation(s) is still far from our nearest marker. If this is the case, then we will have underestimated the phenotypic effect of this QTL, and it alone could explain the presence/absence of escape tunnels. An alternative explanation is that escape-tunnel construction is a threshold trait, but only a high-powered experiment would detect additional loci with either additive or epistatic effects on this behaviour. In either case, any additional QTLs are likely to have small phenotypic effects.

Our results show that QTLs are specific to either entrance-tunnel length or escape-tunnel presence, highlighting the modular nature of the complex *P. polionotus* burrowing behaviour. Moreover, all alleles derived from the *P. polionotus* parent cause an increase in trait value (that is, longer entrance tunnels or the addition of an escape tunnel; Fig. 3b), consistent with the role of natural selection in driving the evolution of the complex burrow architecture. Together, these results show that two aspects of burrow architecture evolved independently, as they are genetically discrete, suggesting that entrance-tunnel length and the presence of an escape tunnel are both ecologically important components of the complex *P. polionotus* burrows.

Although progress has been made towards understanding the genetic underpinnings of innate behaviours, most studies have been restricted to a few laboratory-based model systems[22] or candidate gene approaches in more diverse taxa[23]. By taking advantage of variation in an extended phenotype, in this case burrow architecture in *Peromyscus*, we have gained insight into how a seemingly complex behaviour evolves in the wild. First, we show that the complex, derived burrowing behaviour of *P. polionotus* largely results from evolution at a surprisingly small number of loci. Second, we demonstrate that burrowing behaviour is composed of distinct behavioural modules, which are controlled by independent genetic loci. This is consistent with a century-old observation that elaborate animal behaviour, such as the courtship rituals of grebes, may evolve by integrating simple behaviours[24]. Although genetic modularity has been shown to be important for the evolution of complex morphological traits[25] and laboratory-based behaviours[26], our results extend this pattern to wild

Table 1 | Models for individual and combined QTL effects

Trait	Chromosome	Position	Genotype completeness (%)	lod score	PVE	Additive effect
Log_{10} (average entrance-tunnel length)	1	75.9	77.6	2.48	3.6	0.063
	2	25.1	73.9	3.34	4.9	0.072
	20	59.6	58.8	2.47	3.6	0.064
Full entrance-tunnel-length model	-	-	-	9.55	14.8	-
Escape-tunnel presence (binary)	5	0	60.3	3.38	6.2	0.943

Full model includes all three QTLs. lod, log odds ratio; PVE, percentage of phenotypic variance explained.

behaviours. Thus, we suggest that the behavioural diversity observed in nature may often evolve by accumulating and combining alleles, each with modular effects.

METHODS SUMMARY

Experimental design. We originally obtained outbred stocks of *Peromyscus* from the *Peromyscus* Genetic Stock Center. To start the genetic cross, we mated a single female *P. maniculatus bairdii* with a male *P. polionotus subgriseus* to generate 13 F_1 hybrids, which were then backcrossed to *P. maniculatus* to produce 272 BC-generation hybrids. We assayed burrowing behaviour following our previously described methods[16]. In brief, we placed a single, virgin mouse (60–90 days of age) in an $1.2 \times 1.5 \times 1.1$ m enclosure filled with food, water, nesting material and 700 kg of hard-packed sand under constant temperature and light cycle. Each trial lasted approximately 46 h, and at the end of each trial, we made a polyurethane cast of the resulting burrow[16,27]. From each cast, we measured the length of the entrance tunnel (from the surface to the nest chamber; Fig. 1b) and recorded the presence/absence of an escape tunnel. All mice were assayed in three, consecutive 2-day trials, each time in a new, randomly assigned enclosure.

Genotyping. We extracted DNA using a high-throughput automated phenol-chloroform method (Autogen). We genotyped all individuals using a ddRAD-sequencing approach[20], which identified 526 SNPs with fixed nucleotide differences between the parental species. We used the genotypes of hybrids to estimate genetic linkage among markers. This produced a map containing 24 linkage groups, corresponding to the 24 chromosomes in *P. maniculatus* and *P. polionotus*, with a total map length of 1835.5 cM (Supplementary Fig. 2).

Data analysis. Burrow lengths were log-normalized before correlation and QTL analyses. We performed all statistical tests in the R statistical package[28]. QTLs associated with burrow phenotypes were identified using r/QTL[29]. Specifically, we sequentially performed Haley–Knot regressions, interval mapping, and interval mapping with imputation of missing genotypes, and report QTLs that are consistent across all three methods. We used permutation tests to determine genome-wide statistical significance for each QTL.

Full Methods and any associated references are available in the online version of the paper.

Received 17 August; accepted 27 November 2012.

1. Boake, C. R. B. *et al.* Genetic tools for studying adaptation and the evolution of behavior. *Am. Nat.* **160,** S143–S159 (2002).
2. West-Eberhard, M. J. *Developmental Plasticity and Evolution* (Oxford Univ. Press, 2003).
3. Mackay, T. F. C. The genetic architecture of quantitative traits. *Annu. Rev. Genet.* **35,** 303–339 (2001).
4. Dawkins, R. *The Extended Phenotype* (W. H. Freeman, 1982).
5. Hansell, M. H. *Animal Architecture* (Oxford Univ. Press, 2005).
6. Lorenz, K. Z. The evolution of behaviour. *Sci. Am.* **199,** 67–78 (1958).
7. Peichel, C. L. *et al.* The genetic architecture of divergence between threespine stickleback species. *Nature* **414,** 901–905 (2001).
8. Steiner, C. C., Weber, J. N. & Hoekstra, H. E. Adaptive variation in beach mice caused by two interacting pigmentation genes. *PLoS Biol.* **5,** 1880–1889 (2007).
9. Sumner, F. B. & Karol, J. J. Notes on the burrowing habits of *Peromyscus polionotus. J. Mamm.* **10,** 213–215 (1929).
10. Hayne, D. W. Burrowing habits of *Peromyscus polionotus. J. Mamm.* **17,** 420–421 (1936).
11. Rand, A. L. & Host, P. Mammal notes from Highland County, Florida. *Bull. Am. Mus. Nat. Hist.* **80,** 1–21 (1942).
12. Houtcooper, W. C. Rodent seed supply and burrows of *Peromyscus* in cultivated fields. *Proc. Indiana Acad. Sci.* **81,** 348–389 (1971).
13. Schwartz, C. W. & Schwartz, E. R. *The Wild Mammals of Missouri* (Univ. Missouri Press, 1981).
14. Baker, R. H. in *Biology of Peromyscus (Rodentia)* (ed. King, J. A.) (American Society of Mammalogists, 1968).
15. Wolfe, J. L. & Esher, R. J. Burrowing behaviour of old-field mice (*Peromyscus polionotus*): a laboratory investigation. *Bio. Behav.* **2,** 343–351 (1977).
16. Dawson, W. D., Lake, C. E. & Schumpert, S. S. Inheritance of burrow building in *Peromyscus. Behav. Genet.* **18,** 371–382 (1988).
17. Weber, J. N. & Hoekstra, H. E. The evolution of burrowing behavior in deer mice. *Anim. Behav.* **77,** 603–609 (2009).
18. Tan, K. H. *Soil Sampling, Preparation, and Analysis* (CRC Press, 2005).
19. Wright, S. A mutation of the guinea pig, tending to restore the pentadactyl foot when heterozygous, producing a monstrosity when homozygous. *Genetics* **20,** 84–107 (1935).
20. Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S. & Hoekstra, H. E. Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *PLoS ONE* **7,** e37135 (2012).
21. Beavis, W. D. in *Molecular Dissection of Complex Traits* (ed. Paterson, A. H.) 431–528 (CRC Press, 1998).
22. Bendesky, A. & Bargmann, C. I. Genetic contributions to behaviour at the gene–environment interface. *Nature Rev. Genet.* **12,** 809–820 (2011).
23. Fitzpatrick, M. J. *et al.* Candidate genes for behavioural ecology. *Trends Ecol. Evol.* **20,** 96–104 (2005).
24. Huxley, J. S. The courtship habits of the great crested grebe (*Podiceps cristatus*) with an addition to the theory of sexual selection. *Proc. Zool. Soc. Lond.* **35,** 253–291 (1914).
25. Mallarino, R. *et al.* Two developmental modules establish 3D beak shape variation in Darwin's finches. *Proc. Natl Acad. Sci. USA* **108,** 4057–4062 (2011).
26. Xu, X. *et al.* Modular genetic control of sexually dimorphic behaviors. *Cell* **148,** 596–607 (2012).
27. Felthauser, M. & McInroy, D. Mapping pocket gopher burrow systems with expanding polyurethane foam. *J. Wildl. Manage.* **47,** 555–558 (1983).
28. R Development Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2011).
29. Broman, K. W., Wu, H., Sen, Ś. & Churchill, G. A. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19,** 889–890 (2003).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank D. Brimmer, A. Chiu, A. Goldberg, J. Hopwood, W. Tong, S. Wolff and the Hoekstra laboratory for assistance with behavioural assays and animal husbandry; D. Haig, B. Ölveczky, N. E. Pierce and J. Sanes for helpful discussions; and Harvard's Office of Animal Resources, particularly J. Rocca and M. O'Donnell. We also thank R. Barrett, A. Bendesky, H. Fisher, E. Kay, H. Metz and W. Tong for comments on the manuscript. This research was funded by Chapman Funds for Vertebrate Locomotion to J.N.W., National Science Foundation grant (IOS-0910164) to J.N.W. and H.E.H., and an Arnold and Mabel Beckman Young Investigator Award to H.E.H.

Author Contributions J.N.W. and H.E.H. conceived and designed the experiments. B.K.P. and J.N.W. generated the ddRAD genotypes. J.N.W. performed the behaviour experiments and analysed the genetic and behavioural data. J.N.W. and H.E.H. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to H.E.H. (hoekstra@oeb.harvard.edu).

METHODS

Field observations and soil analysis. We cast and excavated *P. polionotus* burrows at eight sites, distributed broadly across the species range. First, we removed mice from active burrows by inserting plastic tubing into a burrow entrance tunnel while slowly excavating the tunnel (see Supplementary Video for the method conducted in a laboratory environment). The mice generally exited burrows through their escape tunnels when our excavation neared the nest chamber. Once the mice were removed, we filled the remaining burrow cavity with expansive polyurethane filling foam (either from Hilti, or Great Stuff, Dow Chemicals)[27]. After the foam hardened, we excavated the cast and measured the entrance tunnel, the total burrow length, as well as the depth of the lowest point in the nest from the surface (Supplementary Table 1).

In addition to burrow measurements, we collected and analysed samples of undisturbed soil that immediately surrounded the nest chambers of excavated burrows. Before analysing the soil, we sifted each sample through a standardized sieve (2 mm diameter) and then oven dried the remaining soil. We then performed a gravimetric particle analysis on the dried soils (using a hydrometer method[18]) to quantify clay, sand and silt content (Supplementary Table 1).

Statistical analyses of soil and burrow variation. Initially, we $\log_{10}$-transformed all continuous variables to make them normally distributed. For the soil and natural burrow measurements, we used conservative, non-parametric tests to analyse these data because we had small sample sizes. Specifically, we used Spearman tests for analyses of correlation and Kruskal–Wallis tests for differences in variation among groups. For the data collected in the laboratory, we calculated Pearson correlation coefficients among all behavioural variables, and also tested for correlations between burrow measures and soil moisture, animal mass and animal age. We found no significant correlations between total burrow length and moisture, mass or age ($P > 0.05$). Similarly, an analysis of variance (ANOVA) showed no differences in the total length of burrows constructed in different enclosures ($P > 0.05$). We performed all statistical tests using R[28].

Genetic cross. We performed experiments using outbred mice originally obtained from the *Peromyscus* Genetic Stock Center (PGSC). We initially crossed a single female *P. maniculatus bairdii* with a male *P. polionotus subgriseus* to generate 13 first-generation (F_1) hybrids, which we then backcrossed to *P. maniculatus* to produce 272 BC-generation hybrids. We followed standard husbandry protocols used by the PGSC, and Harvard University's Institutional Animal Care and Use Committee approved our experimental methods (protocol 27-09).

Burrowing assays. We housed animals and assayed their burrowing behaviour under controlled laboratory conditions. Specifically, we set the light cycle to 16 h of light and 8 h dark, and room temperature and humidity remained constant at 21–24 °C and 30–37%, respectively. When mice reached an adult age of 60–90 days, we tested their burrowing behaviour in a semi-natural enclosure (described later). Each burrowing trial involved introducing a single mouse to an enclosure at the start of a dark cycle, and then removing the animal approximately 46 h later. After removing the mice, we made casts of burrows by injecting them with polyurethane filling foam (Hilti). We measured the entrance tunnel and total tunnel length of each burrow directly from the casts, and also recorded the presence or absence of an escape tunnel. We conducted three consecutive burrowing trials on each animal, and moved mice to a different enclosure for each trial.

We assayed burrowing behaviour in ten, $1.2 \times 1.5 \times 1.1$ m enclosures, each filled with 700 kg of Premium Playground Sand (Pharmaserv). Before introducing the mice, we contoured the soil surface into three equally sized sections: a flat lower surface; a ~50-cm-high hill at an angle of 60° to horizontal; and a flat upper surface (Fig. 1a). Each enclosure included approximately 5 g of standard rodent food, a 5 cm^2 cotton nestlet (PharmaServ) and a water bottle. At the completion of each trial, we removed all debris from the enclosures, measured soil moisture with a Hydrosense probe (Campbell Scientific), and then thoroughly mixed the sand. We continuously added sand to the enclosures throughout the experiment to maintain constant soil and moisture levels (5–9% and 17–22% water content on the upper and lower surfaces, respectively).

Genotyping. We initially identified 1,700 SNPs, each fixed within but different between *P. maniculatus* and *P. polionotus*, using a modified RAD-tag approach[20]. Briefly, for each individual, we digested ~1 μg of genomic DNA with two restriction endonucleases (100 units of EcoR1-HF and 20 units of Msp1). We ligated the resulting fragments to sequencing adapters containing unique barcodes for each sample. We then pooled ligation products among individuals and isolated fragments in the size range of 280–320 bp using a Pippin Prep electrophoresis platform (Sage BioSciences). Finally, we amplified the remaining fragments using a Phusion High Fidelity PCR Kit (ThermoFisher Scientific) and sequenced the resulting libraries on a Genome Analyzer II (Illumina). For a more complete description of the computational methods used to analyse short-read sequences and to determine genotypes, see ref. 20.

Linkage map construction. We constructed a linkage map in R/qtl[29] using genotypes from 1,700 markers scored in BC-generation individuals. Initially, we calculated the fraction of recombination events and lod score between all marker pairs. We identified 97 markers with genotype information identical to another marker and removed them from the map. We then grouped markers by varying the maximum fraction of recombination events and minimum lod score allowed between markers on the same linkage group. Because the karyotypes of both species are known ($n = 24$ chromosomes), we varied recombination parameters until we recovered a map with 24 linkage groups, each comprising at least 30 markers. This map also contained linkage groups with fewer than 10 markers; we removed the markers in small linkage groups. We ordered the remaining markers by individually scanning linkage groups in overlapping windows of 4–8 markers and minimizing the frequency of recombination events between markers in each window. Next, we removed all markers genotyped in fewer than 100 individuals (Supplementary Fig. 2) as well as all markers with high error rates[30]. Finally, we imputed missing genotypes among all markers that shared identical map locations and then pruned the remaining marker set to reflect an average intermarker distance of ~5–10 cM. The final linkage map contained 526 markers with an average of 390 (s.d. ± 97) genotypes per individual (Supplementary Fig. 3). Finally, we evaluated our genotyping error rate by comparing the likelihood of our marker data, given our estimated linkage map, under different error regimes.

QTL mapping. We performed Haley–Knott regressions and interval mapping analyses sequentially in R/qtl[29] to identify QTLs contributing to burrowing differences. Using permutation tests, we determined the genome-wide significance level for association between markers and phenotypes ($\alpha = 0.05$, $n = 2{,}000$ permutations) as lod = 3.05 and 3.08 for average entrance-tunnel length and binary escape-tunnel presence, respectively[31]. For mapping of escape-tunnel presence, we excluded 29 animals that had unclear escape-tunnel phenotypes (criteria: built an escape tunnel in only one out of three trials and the escape tunnel was ≤4 cm long). lod scores across all linkage groups for both entrance-tunnel length and escape-tunnel presence are shown in Supplementary Fig. 4. With 272 animals in a backcross design, we had 80% power to detect QTLs with effect sizes of >0.3 s.d., assuming the trait is ~70% heritable[32]. We also scanned for pairwise interactions between loci across all linkage groups (using R/qtl[29]), as well as for QTLs that segregated differentially among our BC families due to the outbred nature of our parental stocks (using QTLrel[33,34]). We found no significant evidence for epistasis or for family differences in QTLs. To determine the additive effect of having different genotypes at markers under QTL peaks, we used the fitqtl() and refineqtl() functions in R/qtl[29] to construct models that estimated the percentage of phenotypic variance explained by each QTL (Table 1).

30. Lincoln, S. E. & Lander, E. S. Systematic detection of errors in genetic linkage data. *Genomics* **14,** 604–610 (1992).
31. Doerge, R. W. & Rebai, A. Significance thresholds for QTL interval mapping tests. *Heredity* **76,** 459–464 (1996).
32. Sen, Ś., Satagopan, J., Broman, K. W. & Churchill, G. A. *R/qtlDesign: Inbred Line Cross Experimental Design* (UC San Francisco: Center for Bioinformatics and Molecular Biostatistics, 2006).
33. Cheng, R. *et al.* Genome-wide association studies and the problem of relatedness among advanced intercross lines and other highly recombinant populations. *Genetics* **185,** 1033–1044 (2010).
34. Cheng, R., Abney, M., Palmer, A. A. & Skol, A. D. QTLRel: an R package for genome-wide association studies in which relatedness is a concern. *BMC Genet.* **12,** 66 (2011).