The Genetics of Burrowing Behavior in Peromyscus Mice in the Lab and the Field

Citation

Permanent link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:42029821

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story
The Harvard community has made this article openly available. Please share how this access benefits you. Submit a story.

Accessibility
The genetics of burrowing behavior in *Peromyscus* mice in the lab and the field

A dissertation presented
by
Nicole L. Bedford
to
The Department of Organismic and Evolutionary Biology
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Organismic and Evolutionary Biology

Harvard University
Cambridge, Massachusetts
February 2019
The genetics of burrowing behavior in *Peromyscus* mice in the lab and the field

**Abstract**

Striking variation in animal behavior has repeatedly evolved among closely-related species over short evolutionary timescales. However, the proximate and ultimate mechanisms that give rise to behavioral diversity remain poorly understood. Here, we leverage variation in burrow architecture among deer mice (genus *Peromyscus*), which ranges from short and simple to long and complex. First, we discover that complete, adult-like burrowing behavior emerges earlier in postnatal development in a complex-burrowing species. We then identify a genetic region associated with both precocious burrowing in juveniles and burrow length in adults. Second, we find that the complex-burrowing species digs cooperatively with both same- and opposite-sex partners. However, opposite-sex pairs are more socially cohesive and more likely to engage in simultaneous, coordinated digging, thereby producing longer burrows than same-sex pairs. Third, we design and implement a custom radio frequency identification (RFID) system to non-invasively monitor burrow use in the wild. We find that, in nature, mice frequent the burrows of close genetic relatives and may receive inclusive fitness benefits by maintaining a network of shared burrows. Last, we develop a novel phenotyping platform for the high-throughput quantification of multiple components of burrowing behavior. Using this video-recorded assay, we uncover variation in the temporal dynamics of burrow construction that contribute to interspecific variation in overall burrow structure. We then use a genetic cross between mice with divergent burrow architectures to test for associations between genotype and behavior. By isolating the particular components of burrowing behavior that explain the bulk of heritable variation between species, we can begin to describe the number and types of mutations required to produce, over evolutionary time, a complex behavior from a relatively simple ancestral form. Taken together, these studies clarify the developmental and genetic mechanisms of burrowing behavior in *Peromyscus* and characterize the social and ecological contexts in which burrowing behavior is expressed.
# Contents

0 **Introduction**

1 *Peromyscus* mice as a model for studying natural variation

1.1 Abstract ................................................................. 6
1.2 Introduction ............................................................ 7
1.3 Distribution and habitat .............................................. 9
1.4 Diet and predators .................................................... 12
1.5 Parasites and disease ............................................... 13
1.6 Longevity ............................................................... 13
1.7 Life history ............................................................. 14
1.8 Mating system and parental care ................................ 15
1.9 Home building ......................................................... 16
1.10 Pigmentation .......................................................... 17
1.11 Adaptation to mountains, deserts and cities ................. 19
1.12 *Peromyscus* and the history of evolutionary thought .... 19
1.13 *Peromyscus* in the Laboratory .................................. 21
1.14 Conclusions ............................................................ 22
1.15 Box 1. Future directions for *Peromyscus* research ........ 23
1.16 References ............................................................. 24

2 Evolution and genetics of precocious burrowing behavior in *Peromyscus* mice

2.1 Abstract ................................................................. 32
2.2 Results ................................................................. 33
2.3 Discussion .............................................................. 34
2.4 References ............................................................. 34

3 Behavioural mechanisms of cooperative burrowing in *Peromyscus* mice

3.1 Abstract ................................................................. 50
3.2 Introduction ............................................................ 51
3.3 Results ................................................................. 51
### Listing of figures

1.1 Simplified phylogeny depicting the relationships among muroid rodent model organisms. .......................................................... 8  
1.2 North American distributions of eight *Peromyscus* species. .......................................................... 9  
1.3 The ecology of *Peromyscus* varies both within and among species. .......................................................... 12  
1.4 Genetic crosses between the pale beach mouse *P. polionotus leucocephalus* and the darker mainland mouse *P. p. polionotus*. .......................................................... 18  
2.1 The ontogeny of burrow construction in two sister species of *Peromyscus*. .......................................................... 36  
2.2 Reciprocal interspecific cross-fostering. .......................................................... 38  
2.3 Genetic dissection of precocious burrowing in *P. polionotus* × *P. maniculatus* hybrids. .......................................................... 40  
2.4 Effect of the dominant *P. polionotus* allele on linkage group 2 on burrowing behavior in backcross hybrids. .......................................................... 42  
3.1 *P. polionotus* pairs cooperatively construct burrows that are nearly twice as long as those dug by individuals. .......................................................... 55  
3.2 *P. polionotus* females dig longer burrows after cohabitation with a male, regardless of pregnancy status. .......................................................... 57  
3.3 Sex differences in digging behaviour across social contexts. .......................................................... 59  
3.4 Opposite-sex pairs have more affiliative and fewer agonistic interactions. .......................................................... 62  
3.5 Opposite-sex pairs are more likely to engage in efficient simultaneous digging. .......................................................... 63  
4.1 Home-range characterization and burrow use patterns as determined by capture-recapture and RFID-monitoring. .......................................................... 75  
4.2 Groups of mice visit networks of spatially clustered burrows. .......................................................... 77  
4.3 Nightly activity summaries, distance between burrows, and burrow use evenness. .......................................................... 78  
4.4 Distribution of RFID timestamps and fraction of spatiotemporal overlap in the population as a function of time-bin size. .......................................................... 78  
4.5 Spatiotemporal overlap at burrow entrances. .......................................................... 79  
4.6 Genetic relatedness predicts spatiotemporal overlap between pairs. .......................................................... 81  
4.7 Close relatives sleep in the same burrow. .......................................................... 81  
4.8 Distributions of burrow site attributes for 36 RFID-monitored burrows. .......................................................... 82
4.9 Long burrows are better buffered against severe temperature fluctuations.

5.1 Novel burrowing assay reveals key differences in burrowing behavior between two Peromyscus species.

5.2 Trait correlations reveal additional species differences in burrowing behavior.

5.3 A QTL for time spent digging.

5.4 QTLs for component traits of burrowing behavior.

A.1 Morphological and motor development in P. polionotus and P. maniculatus.

A.2 Effect of genotype on juvenile and adult burrowing behavior in hybrids.

B.1 Different pair configurations differ in total time spent digging, burrow length, and division of labour.

C.1 Beach mouse home-range sizes estimated from utilization distributions.

C.2 Home-range characterization and burrow use patterns as determined by capture-recapture and RFID-monitoring.

C.3 Groups of mice visit networks of spatially clustered burrows.

C.4 Nightly activity summaries, distance between burrows, and burrow use evenness.

C.5 Distribution of RFID timestamps and fraction of spatiotemporal overlap in the population as a function of time-bin size.

C.6 Spatiotemporal overlap at burrow entrances.

C.7 Close relatives show greater spatiotemporal overlap.

D.1 Pearson correlation coefficients for pairs of component burrowing traits.
I would like to thank Hopi and members of the Hoekstra Lab, both past and present, who have been such supportive friends and colleagues. Many thanks are owed to my Dissertation Committee for sharing their expertise and advice. I would also like to thank members of the OEB community, particularly my cohort who have been sources of inspiration and solidarity throughout the various stages of graduate school. I thank my family for always encouraging me to do my best. Finally, I thank Adam Nelson who, from the very beginning, has provided love and support in countless ways.
Introduction
In 1963, Niko Tinbergen outlined a set of four guiding questions that would unify and focus the various disciplines that, at the time, fell loosely under the emerging science of Ethology. Collectively, the four questions—ontogeny, evolution, survival value, and causation—address both historical and contemporary processes as well as proximate and ultimate causes. In principle, this approach could be applied to any problem in Biology, but it has proven particularly useful for clarifying thinking about behavior. Tinbergen argued that “a comprehensive, coherent science of Ethology has to give equal attention to each of them and to their integration.” Under this framework, I present a series of studies that, taken together, offer a broad understanding of burrowing behavior in North American deer mice (genus *Peromyscus*).

Burrowing in *Peromyscus* presents a unique opportunity to advance a complete ethological account of behavior, due to the variability, utility, and tractability of the trait. First, *Peromyscus* mice produce a variety of stereotyped, species-specific burrow architectures. Second, these variable structures correlate with habitat: species that construct long burrows tend to occupy environments with poor vegetative cover such as abandoned agricultural fields, whereas species that construct short burrows occupy more structured environments such as grasslands or forests. This observation suggests a special utility for long burrows; namely providing refuge from predation and severe climate in otherwise exposed habitats. Interspecific variation in burrow architecture, therefore, is likely to be adaptive. Last, burrowing behavior and burrow structures are easily quantified in a controlled laboratory setting. Remarkably, mice raised in the lab for several generations recapitulate the species-specific burrow architectures observed in nature. This suggests that variation in burrowing behavior has a strong genetic component, and that the underlying genetic mechanisms are tractable. Together, these attributes qualify burrowing behavior as a prime candidate for comprehensive examination according to Tinbergen’s framework.

In Chapter 1, I review the emergence of *Peromyscus* as a model organism and the wealth of natural history data available for the genus. Deer mice have adapted to a wide array of ecological condi-
tions across North America and, as a result, are excellent models for the study of natural variation in mammals. This chapter examines the vast body of work on *Peromyscus* biology and lays the foundation for subsequent chapters.

In Chapter 2, I investigate the ontogeny of burrowing behavior in two *Peromyscus* sister-species. By measuring the developmental progression of burrowing behavior in a backcross population, we identify a genetic region that affects different aspects of burrowing behavior during different life history stages. Variation in this genetic region predicts both the age at which juvenile mice first construct a burrow as well as the length of the burrow in adulthood. This suggests that a single gene—or multiple linked genes—alters the motivation to burrow at multiple timepoints in development.

In Chapter 3, I take a comparative, experimental approach to understand the role of social context in the evolution of behavioral plasticity. In a monogamous species—but not two closely-related promiscuous species—we find that opposite-sex pairs construct longer burrows than same-sex pairs. Through statistical modelling, we show that opposite-sex pairs are more likely to dig in a coordinated fashion in the same burrow—a more efficient mode of burrow elongation that produces longer burrows despite no change in total digging duration by either member of the pair. This indicates that plasticity in burrow structure can be achieved without pronounced changes in individual burrowing effort.

In Chapter 4, I design a custom RFID-tracking system to document patterns of burrow use in a wild population. We find that, in addition to providing a retreat from predators and severe climate, burrows are arenas for complex social dynamics. In the wild, a mouse will visit a network of several, spatially-clustered burrows during the night. Mice tend to visit the same networks as close relatives, which suggests inclusive fitness benefits to maintaining a shared network of actively-used burrows. Accordingly, I argue that burrows are adaptations that improve the fit between organism and environment.

In Chapter 5, we develop a novel assay to decompose variation in burrowing behavior between
two species with strikingly different burrow morphologies (i.e., simple and complex). We find that the complex burrower spends more time digging, and digs more efficiently, than the simple burrower. Furthermore, variation in different components of burrowing behavior is associated with variation in multiple genomic regions. Using quantitative trait locus (QTL) mapping, we show that digging behaviors map to a different genetic region than other movement-related behaviors, indicating that complex behavior can be broken down into its constituent parts to reveal distinct genetic underpinnings for each component.

Taken together, these studies, and the work they draw upon, address various aspects of the ontogeny, evolution, survival value, and causation of burrowing behavior in *Peromyscus*. This multifaceted approach offers new insights into a complex trait and provides important context for interpreting the nature of the genetic changes that ultimately lead to the evolution of behavioral diversity.
References

Within the range of one species (maniculatus) it is probable that a line, or several lines, could be drawn from Labrador to Alaska and thence to southern Mexico throughout which not a single square mile is not inhabited by some form of this species.

Wilfred H. Osgood

1

*Peromyscus* mice as a model for studying natural variation
1.1 Abstract

Deer mice (genus *Peromyscus*) are the most abundant mammal in North America and, throughout their wide range, occupy nearly every terrestrial habitat type. It may thus be unsurprising that the natural history of *Peromyscus* is among the best studied of any small mammal. For decades, the deer mouse has contributed to our understanding of population genetics, disease ecology, longevity, endocrinology and behaviour. Over a century’s worth of detailed descriptive studies of *Peromyscus* in the wild, coupled with emerging genetic and genomic techniques, have now positioned these mice as model organisms for the study of natural variation and adaptation. Recent work, combining field observations and laboratory experiments, has lead to exciting advances in a number of fields—from evolution and genetics, to physiology and neurobiology.

1.2 Introduction

*Peromyscus* is a genus of small North American rodents, known colloquially as deer mice (Emmons 1840). When the first *Peromyscus* specimens were shipped to European systematicists in the late 18th century, their resemblance to the local wood mouse prompted the designation *Mus sylvaticus* (Hooper 1968). At the time, little was known of the diversity of rodents worldwide and most were assigned the generic term *Mus* (Linnaeus 1758). The name *Peromyscus* (Gloger 1841) was first employed, albeit narrowly, mid 19th century. Quadrupeds of North America (Audubon and Bachman 1851-1854) recognized only three species now known to belong to *Peromyscus*, and Mammals of North America (Baird 1859) included a mere twelve. But by the turn of the 20th century, *Peromyscus* included 143 forms, 42 of which represented monotypic or good biological species (Osgood 1909).
The genus saw several additional revisions throughout the 20th century as North American mammalogy matured and natural history collections expanded. Today 56 species are recognized, the most widespread and diverse being *P. maniculatus* (Musser and Carleton 2005).

Thus, although not immediately appreciated, *Peromyscus* includes more species than any other North American mammalian genus and, apart from *Mus* and *Rattus*, more is known concerning its biology in the laboratory than any other group of small mammals (Fig. 1; King 1968, Kirkland and Layne 1989). Several disciplines including ecology, evolution, physiology, reproductive biology and behavioural neuroscience have all employed *Peromyscus*, inspiring its label as “the Drosophila of North American mammalogy” (Dewey and Dawson 2001). Arguably, the emergence of *Peromyscus* as a model system was propelled by our cumulative knowledge of its fascinating and varied natural history.

![Figure 1.1: Simplified phylogeny depicting the relationships among muroid rodent model organisms. *Peromyscus* belong to the Cricetidae family, which includes voles (*Microtus*), hamsters (*Mesocricetus*), and New World rats and mice. Old World rats and mice belong to the Muridae family, which include the familiar laboratory rat (*Rattus norvegicus*) and mouse (*Mus musculus*). Muridae and Cricetidae diverged roughly 25 million years ago (redrawn from Steppan et al. 2004).](image-url)
Figure 1.2: North American distributions of eight *Peromyscus* species currently maintained as outbred laboratory stocks at the Peromyscus Genetic Stock Center and various universities (redrawn from Hall 1981). Simplified tree indicating phylogenetic relationships among taxa is shown; branch lengths are arbitrary (redrawn from Bradley et al. 2007). The most widespread and ecologically diverse group is also the best represented in the laboratory: six *P. maniculatus* subspecies are maintained in laboratories across the United States. Collecting localities of all colony founders are indicated by numbered squares (see Table 2).

1.3 Distribution and habitat

Osgood (1909) asserted that some form of *Peromyscus* had been trapped in nearly every patch of North America ever visited by a mammal collector. Members of the genus are distributed from the southern edge of the Canadian Arctic to the Colombian border of Panama (Fig. 2). Various demo-
Table 1.1: Notable museum collections of Peromyscus skins, skulls, and skeletons.

<table>
<thead>
<tr>
<th>Collection</th>
<th>Location</th>
<th>No. Specimens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smithsonian National Museum of Natural History</td>
<td>Washington, DC</td>
<td>38,406</td>
</tr>
<tr>
<td>Museum of Vertebrate Zoology</td>
<td>Berkeley, CA</td>
<td>34,131</td>
</tr>
<tr>
<td>American Museum of Natural History</td>
<td>New York, NY</td>
<td>19,234</td>
</tr>
<tr>
<td>Field Museum</td>
<td>Chicago, IL</td>
<td>8,939</td>
</tr>
<tr>
<td>Museum of Comparative Zoology</td>
<td>Cambridge, MA</td>
<td>7,754</td>
</tr>
<tr>
<td>Canadian Museum of Nature</td>
<td>Ottawa, ON</td>
<td>6,315</td>
</tr>
<tr>
<td>Academy of Natural Science</td>
<td>Philadelphia, PA</td>
<td>2,425</td>
</tr>
<tr>
<td>Natural History Museum</td>
<td>London, UK</td>
<td>2,238</td>
</tr>
</tbody>
</table>

graphic and biogeographic factors (e.g. Pleistocene glacial and pluvial cycles, population expansions, mountain range elevations and sea-level changes) have influenced the diversity and distribution of deer mice (Sullivan et al. 1997, Riddle 2000, Dragoo et al. 2006, Kalkvik et al. 2012, López-González et al. 2014). The result is a mosaic of widespread and restricted species ranges shaped by both dispersal and vicariance events. Our knowledge of the distributions, home ranges and habitat preferences of deer mice comes primarily from the trapping data and field notes of early natural historians (e.g. Sumner 1917, Dice 1931, Blair 1940, 1951). Osgood’s influential 1909 taxonomic revision was built on examinations of more than 27,000 specimens from diverse locales that were collected primarily by the U.S. Biological Survey. Today, more than 120,000 Peromyscus specimens are accessioned in Natural History museums across North America and the United Kingdom (Table 1). These invaluable collections document more than a century of dynamic relationships between deer mice and their environment. For example, by comparing past and present-day collecting locales, shifts in the distributions of deer mice have been linked to climate change (Moritz et al. 2008, Yang et al. 2011, Rowe et al. 2014), and morphological analyses of these museum specimens reveal how deer mice might respond to changing environments (Grieco and Rizk 2010).

Although not strictly commensal, deer mice (particularly in New England) do enter human households and partake of their larders. According to legend, Walt Disney drew inspiration for
Table 1.2: Peromyscus laboratory colonies currently employed in research (as of June 2015). The year and population from which the founders were collected are noted. PGSC = Peromyscus Genetic Stock Center, UNM = University of New Mexico, HU = Harvard University, UIUC = University of Illinois at Urbana-Champaign.

<table>
<thead>
<tr>
<th>No.</th>
<th>Species</th>
<th>Year</th>
<th>Source Population</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><em>P. californicus insignis</em></td>
<td>1979-1987</td>
<td>Santa Monica Mts., CA</td>
<td>PGSC</td>
</tr>
<tr>
<td>2</td>
<td><em>P. eremicus sp.</em></td>
<td>1993</td>
<td>Tucson, AZ</td>
<td>PGSC</td>
</tr>
<tr>
<td>3</td>
<td><em>P. polionotus subgriseus</em></td>
<td>1952</td>
<td>Ocala National Forest, FL</td>
<td>PGSC</td>
</tr>
<tr>
<td>4a</td>
<td><em>P. maniculatus bairdii</em></td>
<td>1946-1948</td>
<td>Ann Arbor, MI</td>
<td>PGSC</td>
</tr>
<tr>
<td>4b</td>
<td><em>P. m. sonoriensis</em></td>
<td>1995</td>
<td>White Mtn. Research Station, CA</td>
<td>PGSC</td>
</tr>
<tr>
<td>4c</td>
<td><em>P. m. rufinus</em></td>
<td>1998</td>
<td>Manzano Mtn., NM</td>
<td>UNM</td>
</tr>
<tr>
<td>4d</td>
<td><em>P. m. nubiterrae</em></td>
<td>2010</td>
<td>Powder Mill Nature Reserve, PA</td>
<td>HU</td>
</tr>
<tr>
<td>4e</td>
<td><em>P. m. rufinus</em></td>
<td>2014</td>
<td>Mt. Evans, CO</td>
<td>UIUC</td>
</tr>
<tr>
<td>4f</td>
<td><em>P. m. nebrascensis</em></td>
<td>2014</td>
<td>Lincoln, NE</td>
<td>UIUC</td>
</tr>
<tr>
<td>5</td>
<td><em>P. leucopus sp.</em></td>
<td>1982-1985</td>
<td>Linville, NC</td>
<td>PGSC</td>
</tr>
<tr>
<td>6</td>
<td><em>P. gossypinus gossypinus</em></td>
<td>2009</td>
<td>Jackson County, FL</td>
<td>HU</td>
</tr>
<tr>
<td>7</td>
<td><em>P. melanophrys xenerus</em></td>
<td>1970-1978</td>
<td>Zacatecas, Mexico</td>
<td>UIUC</td>
</tr>
<tr>
<td>8</td>
<td><em>P. azteca hylocetes</em></td>
<td>1986</td>
<td>Sierra Chincua, Mexico</td>
<td>UIUC</td>
</tr>
</tbody>
</table>

Mickey Mouse from the “tame field mice” (most likely *P. leucopus*) that would wander into his old Kansas City animation studio (Updike 1991). However, *Peromyscus* are most commonly trapped in woodlands and brushlands, and are also found in tropical and temperate rainforests, grasslands, savannas, swamps, deserts and alpine habitats (Fig. 3; Baker 1968). Local adaptation to these various environments has been the subject of much recent inquiry (e.g. Linnen et al. 2013, Natarajan et al. 2013, MacManes and Eisen 2014), and the detailed cataloguing of phenotypic diversity by early naturalists inspired much of this work. However, we still require a more complete understanding of ecological diversity across the entire genus, as well as an enlightened view of phylogenetic relationships informed by whole-genome sequences (Box 1).
Figure 1.3: The ecology of *Peromyscus* varies tremendously both within and among species. A) The forest-dwelling deer mouse, *P. maniculatus nubiterrae*, perches high on a tree branch in Southwestern Pennsylvania. B) The beach mouse, *P. polionotus phasma*, takes shelter among the dune grasses on Florida’s Atlantic coast. C) Its mainland counterpart, the oldfield mouse, *P. polionotus sumneri*, is typically found in fallow fields and is sympatric with D) the cotton mouse, *P. gossypinus*, which occupies adjacent stands of long leaf pine. Both species (C and D) were trapped at Lafayette Creek Wildlife Management Area in Walton County, Florida. Image credits: A, E. P. Kingsley; B, J. B. Miller; C, D, N. L. Bedford.

### 1.4 Diet and predators

Generally deer mice are granivores, feeding primarily on seeds, but fruits, fungi, green vegetation and insects have been found among their stomach contents and in the nest cavities of their burrows (Gentry and Smith 1968, Wolff et al. 1985). However, some species have evolved seasonally specialized diets. In the winter, *P. melanotis* prey almost exclusively on monarch butterflies that roost in Mexico’s central highlands (Brower et al. 1985). Moreover, on a remote island in British Columbia, *P. keeni* feast on auklet eggs during the seabird breeding season (Drever et al. 2000). Deer mice are themselves common prey, contributing to the diets of many predators such as weasels, skunks, lynx, bobcats, foxes, coyotes, hawks and owls (Luttich et al. 1970, Bowen 1980, Montgomery 1989, Van Zant and Wooten 2002). Indeed, avian predation imposes strong selective pressure for cryptic
colouration in *Peromyscus*—a classic example of local adaptation (Vignieri et al. 2010, Linnen et al. 2013).

1.5 Parasites and disease

The diversity of parasites is documented for only a few *Peromyscus* species, and very little is known of the ecological factors that influence infection dynamics. Common internal parasites include pentastomid larvae, cestode tapeworms, nematodes and trematodes (Whitaker 1968, Pederson and Antonovics 2013). External parasites include lice, mites, fleas and ticks (Whitaker 1968), the latter two being vectors of plague and Lyme disease, respectively (Allred 1952, Burgdorfer et al. 1982, Gage and Kosoy 2005).

As a natural reservoir for *Borrelia burgdorferi*—the bacterial agent of Lyme disease—*Peromyscus* is a feature of much research on the pathogenesis and transmission of the disease (Bunikis et al. 2004, Ramamoorthi et al. 2005, Schwanz et al. 2011, Baum et al. 2012). *Peromyscus* also features in ecological modeling efforts to determine how the diversity of the tick host community impacts disease risk (LoGiudice et al. 2003, 2008). One hypothesis for the alarming recent expansion of Lyme disease is that habitat fragmentation associated with human development favours deer mouse populations at the expense of other tick hosts (e.g. squirrels and shrews) that are poor reservoirs for the disease (LoGiudice et al. 2003, Schwanz et al. 2011). *Peromyscus* is also a notorious carrier of the Sin Nombre hantavirus, responsible for the deaths of 12 people in the Four Corners area of the southwestern United States in 1993.

1.6 Longevity

Mortality in natural populations is incredibly high and driven by a combination of factors including limited food supply, competition for territories and predation (Bendell 1959). As such, most
Peromyscus are thought to live less than a year in the wild (Terman 1968). However, early investigators noted substantially longer natural lifespans in their laboratory colonies (Sumner 1922, Dice 1933). With a two-fold difference in life expectancy, Sacher and Hart (1978) proposed *P. leucopus* and *Mus musculus* as a longevity contrast pair. *P. leucopus*—which lives up to 8 years and remains fertile for 5—produces fewer reactive oxygen species, exhibits enhanced antioxidant enzyme activity and less oxidative damage to lipids relative to the short-lived (3.5 years) laboratory mouse (Sohal et al. 1993, Shi et al. 2013). Measuring the biochemical correlates of longevity in *Peromyscus* has been integral to providing support for the oxidative stress theory of aging (Ungvari et al. 2008).

1.7 Life history

The timing of life history events in *Peromyscus*—well documented from field and laboratory studies alike—is highly variable both within and among species. Yet studies contrasting the reproductive and developmental patterns of wild and domesticated deer mice have found few significant differences (Millar 1989, Botten et al. 2001). Here, we highlight life history traits in *P. maniculatus*, the most commonly used laboratory species. Gestation ranges from 21 to 27 days (average 23.6) and average litter size is 4.6 pups (Millar 1989). Juveniles first leave the nest between 14 and 16 days of age (Vestal et al. 1980) and become independent of their mother between 18 and 25 days (Millar 1989). Captive females give birth to their first litter, on average, at 84 days (Haigh 1983), but males are capable of siring offspring several weeks earlier.

The actual timing of sexual maturation in the wild, however, is often dictated by population density, food availability and season. In response to short day length, many species exhibit seasonal gonadal regression (Trainor et al. 2006), increased aggression (Trainor et al. 2007), impaired spatial memory (Workman et al. 2009) and enhanced immune function (Prendergast and Nelson 2001). As such, *Peromyscus* has emerged as a model system for the study of photoperiodism (i.e. the ability
to seasonally modulate energetic demands by tracking day length changes). Such studies have been particularly fruitful for understanding the mechanistic basis of gene by environment interactions. For example, day length can reverse the behavioural action of the hormone estradiol by determining which estrogen receptor pathway is expressed and consequently activated (Trainor et al. 2007). While life history traits are strongly affected by environmental cues, substantial genetic variation in the neuroendocrine pathways that control reproductive timing also exists, as demonstrated by selection line experiments with photoperiod responsive and nonresponsive *P. leucopus* (Heideman et al. 1999, Heideman and Pittman 2009).

1.8 *Mating system and parental care*

While the majority of *Peromyscus* species are promiscuous, monogamy has independently evolved at least twice in the genus (Turner et al. 2010). Both *P. californicus* (Gubernick and Alberts 1987, Ribble 1991) and *P. polionotus* (Smith 1966, Foltz 1981) are socially and genetically monogamous, and both males and females contribute to the care of offspring. *Peromyscus* californicus, in particular, has become an important neurobiological model for the study of male parental care (Bester-Meredith et al. 1999, Trainor et al. 2003, Lee and Brown 2007, de Jong et al. 2009). As a complement, the ability of monogamous *P. polionotus* to hybridize with promiscuous *P. maniculatus* allows geneticists to identify the genetic basis of alternate mating systems and their associated phenotypes, from genomic imprinting (Vrana et al. 2000) to parental investment and reproductive traits (e.g. Fisher and Hoekstra 2010).

Rosenfeld (2015) argues that parental and social behaviours are particularly vulnerable to endocrine disruption, as these traits are dependent upon the organizational and activational effects of androgens and estrogens. Mating system variation between closely related species of deer mice provides an opportunity to test this hypothesis. *Peromyscus* maniculatus males exposed to the endocrine
disrupting compound bisphenol A (BPA) during development displayed reduced spatial learning and exploratory behaviour—traits known to be associated with male-male competition for mates (Galea et al. 1996, Jašarević et al. 2011). However, these behaviours—which are not subject to sexual selection in females—were unaffected in BPA-exposed females. By contrast, sexual selection favours the evolution of mate guarding and territorial behaviour in monogamous males, and it is these traits (rather than spatial learning or exploratory behaviour) that are compromised by endocrine disruption in *P. californicus* (Williams et al. 2013).

### 1.9 Home building

Behavioural genetics studies have historically been restricted to a handful of genetic model organisms that display behaviours of unclear ecological relevance (Fitzpatrick et al. 2005). The structures built by animals (e.g. termite mounds, bird nests, rodent burrows) are promising candidates for the genetic dissection of natural behaviour because these "extended phenotypes" can be precisely measured and because, as extended products of the actions of genes, they are subject to natural selection (Dawkins 1982). Sufficient resources are now available—from a medium-density genetic linkage map (Kenney-Hunt et al. 2014) to draft genome sequences (Baylor College of Medicine, *Peromyscus* Genome Project)—that we can attribute natural variation in *Peromyscus* behaviour to specific genetic variants.

Recently diverged sister-species *P. maniculatus* and *P. polionotus* (Avise et al. 1983), display considerable differences in stereotyped burrowing behaviour. *Peromyscus* maniculatus digs short, simple burrows in contrast to the long, complex burrows constructed by *P. polionotus* that consist of an entrance tunnel, nest chamber and escape tunnel (Dawson et al. 1988, Weber et al. 2013). Remarkably, mice raised in the laboratory for several generations recapitulate the species-specific burrow architectures observed in nature (Video 1). Furthermore, the complex burrows of *P. polionotus* are
derived (Weber and Hoekstra 2009) and likely evolved through changes at only a handful of genetic loci, each affecting distinct behavioural modules (i.e. entrance tunnel length and escape tunnel presence; Weber et al. 2013). While the next steps include isolating genetic variants and understanding their effects on the neural circuitry underlying burrowing behaviour, the full adaptive significance of burrowing in the wild remains unclear.

1.10 Pigmentation

Among the several cases of adaptive phenotypic variation in Peromyscus, perhaps the most obvious is coat colouration. Recent advances have identified not only the genes, but also the specific mutations leading to local variation in coat colour. Beach mice (P. polionotus leucocephalus) living on the coastal sand dunes and barrier islands of Florida are considerably paler than their inland counterparts (P. p. subgriseus) that inhabit dark, loamy soils (Fig. 4; Howell 1920, Sumner 1929). For beach mice on Florida’s Gulf Coast, light colouration is due, in part, to a fixed single nucleotide polymorphism (SNP) in the melanocortin-1 receptor (Mc1r) coding region (Hoekstra et al. 2006). However, this Mc1r allele does not contribute to light pelage in Florida’s Atlantic coast mice, suggesting that the two populations converged on light colouration independently (Steiner et al. 2007).

Similarly, background matching in P. maniculatus of the Nebraska Sand Hills affords a strong
selective advantage against avian predators (Linnen et al. 2013). Yet, cryptic colouration is a complex phenotype composed of multiple component traits (i.e. tail stripe, dorsal-ventral boundary, ventral colour, dorsal brightness and hue). Linnen and colleagues (2013) identified multiple distinct mutations within the Agouti locus, each associated with a different colour trait that independently affected fitness. Thus, parallel studies of *Peromyscus* pigmentation nicely illustrate the marriage between classical natural history studies and modern molecular techniques, thereby providing new insights into the molecular basis of adaptation.

![Image of mice with varying colors](image)

**Figure 1.4:** Genetic crosses between the pale beach mouse *P. polionotus leucocephalus* (top row, left) and the darker mainland mouse *P. p. polionotus* (top row, right) result in first-generation F1 hybrids, all with intermediate colouration (2nd row, center). Intercrosses between F1 hybrids produce a variable F2 generation showing a continuous distribution of pigmentation phenotypes ranging from light to dark (3rd and 4th rows). This segregation pattern—initially described by Francis Sumner—is among the earliest empirical evidence that several discrete loci may collectively contribute to a quantitative trait (Dobzhansky 1937). *Peromyscus* thus played a key role in reconciling Mendelian heredity with the Darwinian view that evolution proceeds by the gradual accumulation of small variations. Image credit: N. L. Bedford.
Among North American mammals, the deer mouse is unparalleled in its ability to colonize an impressive array of habitats. The remarkable elevational range of one subspecies (*P. m. sonoriensis*) stretches from below sea level in Death Valley to above 4,300 meters in the adjacent White and Sierra Nevada mountain ranges (Hock 1964). The ability of deer mice to colonize and thrive in low-oxygen environments is due, in part, to standing genetic variation in globin genes (Snyder 1981, Natarajan et al. 2015). Storz and colleagues (2007, 2009) pinpointed several amino acid substitutions that confer high hemoglobin-O2 affinity and better aerobic performance at high altitudes. Functional analyses have since identified the precise effects of each mutation, and interactions among mutations, on hemoglobin-O2 affinity, demonstrating that the adaptive value of a given biochemical substitution depends both on the local environment and the genetic background in which it arises (Natarajan et al. 2013). Urbanization, as well, can influence the genetic structure of *Peromyscus* populations (Pergams and Lacy 2008, Munshi-South and Kharchenko 2010, Munshi-South and Nagy 2014). Moreover, Harris et al. (2013) compared brain, liver and gonad transcriptomes for urban and rural populations of *P. leucopus*, thereby identifying several genes associated with metabolism and immune function that exhibit signatures of selection in New York City parklands. Similarly, MacManes and Eisen (2014) identified renal transcripts related to solute and water balance experiencing purifying selection in the desert-adapted species, *P. eremicus*. Further study of these candidate genes will determine their role in adaptation to extreme or new environments.

**Peromyscus and the history of evolutionary thought**

The work of early *Peromyscus* biologists (particularly Francis B. Sumner) informed influential thinkers in population genetics and evolutionary biology such as Sewall Wright, Theodosius Dobzhansky and J.B.S. Haldane. Since most early 20th century geneticists came from experimental-
ist backgrounds, many turned to naturalists for data from wild populations (Provine 1986). At the time, Sumner’s work on geographic variation in *Peromyscus* represented one of the few major studies of evolution in natural populations. As such, Wright closely followed Sumner’s analysis of phenotypic intergradation between geographically contiguous *P. maniculatus* subspecies in California (Sumner 1918). Wright concluded that the observed quantitative differences in coat color were determined by the accumulation of several discrete (i.e. Mendelian) factors (Wright 1932). The question of whether continuous (or quantitative) traits are subject to the same rules of inheritance as discrete characters was central to the Modern Evolutionary Synthesis.

Between 1914 and 1930 Sumner made careful measurements of several quantitative traits—most notably coat colour—that varied among geographically distinct subspecies of *Peromyscus*, which he then crossed in the laboratory (Sumner 1930). Dobzhansky (1937) highlighted these data as empirical support for the multiple gene hypothesis for the inheritance of quantitative traits (Fig. 4). Later, in his 1948 paper *The Theory of a Cline*, Haldane applied a theoretical model to the gradient of increasing pigmentation observed in *P. polionotus* populations as one moves inland from coastal Florida (Sumner 1929). Haldane used this intergradation to estimate the local strength of selection acting on a putative pigmentation locus in the wild—the dominant white-cheek character (Wc) identified by Blair (1944).

Deer mice also featured prominently in Dobzhansky’s consideration of the mechanisms of reproductive isolation. Certain *P. maniculatus* subspecies with overlapping geographic distributions are nevertheless separated by habitat, often with one subspecies inhabiting prairie, open fields or sandy lake beaches, and the other being exclusively forest-dwelling (Dice 1931). These subspecific forms readily produce viable and fertile offspring in the laboratory yet remain reproductively isolated in the wild—a prime example of ecological isolation (Dobzhansky 1937). *Peromyscus* has thus been a cornerstone of evolutionary biology for nearly a century. These studies and others drew the attention of biologists in many fields; launching the multifold *Peromyscus* research programs we see
today.

1.13 *Peromyscus* in the Laboratory

Sumner, considered the grandfather of *Peromyscus* biology, first demonstrated the feasibility of the deer mouse as a laboratory organism in the 1910s and 20s. He famously built the first *Peromyscus* “mouse house” in what is now referred to as Sumner Canyon at the Scripps Institution in La Jolla, California. When his *Peromyscus* work at Scripps was discontinued, Sumner bequeathed his stocks to Lee R. Dice at the University of Michigan who honed the methods for generating and maintaining *Peromyscus* colonies in the 1930s and 40s. During this time, Dice began to catalogue single factor genetic mutations in his stocks (e.g. gray, dilute, epilepsy). These mice served as the founding strains for the *Peromyscus* Genetic Stock Center (PGSC), established in 1985 by Wallace Dawson at the University of South Carolina, which currently maintains wild-derived stocks of 6 species as well as 13 coat-colour mutants and 4 additional mutants on *P. maniculatus* genetic backgrounds. Additional wild-derived stocks are kept in individual laboratories (Table 2) and still more mutants have been cryopreserved. The PGSC also maintains an extensive online reference library (http://stkctr.biol.sc.edu) with more than 3000 citations.

While the genetic causes and phenotypic consequences differ among strains, *Peromyscus* colonies are invariably susceptible to inbreeding depression, which necessitates their maintenance as relatively outbred stocks (Lacy et al. 1996, Joyner et al. 1998). Thus, although the deer mouse is amenable to laboratory life, its biology has not been purposely altered by generations of inbreeding or artificial selection. Life history traits and even behaviours such as burrow construction or ultrasonic vocalization are generally preserved in laboratory strains (Millar 1989, Dawson et al. 1988, Kalcounis-Rueppell et al. 2010). Thus, the traits we scrutinize in the laboratory (e.g. aerobic performance, photoperiodism, mating and parental behaviour) are arguably faithful representations of phenotypes
in nature. The ability to study genetically diverse, wild-derived mice under controlled laboratory conditions has opened up several constructive research programs centered on understanding the phenotypic consequences of natural genetic variation.

1.14 Conclusions

The tradition of dissecting the genetic basis of ecologically relevant traits in the laboratory began in the early 20th century; in *Peromyscus*, this effort was lead by Francis Sumner and continues today. In an era of high-throughput sequencing and expanding transgenic technologies, our concept of the genetic model organism is rapidly changing. We can now widen our focus to include the diverse and naturally evolving species that may further our understanding of life outside the laboratory. The emergence of *Peromyscus* as a model system has been largely driven by the wealth of natural history information available for the genus. Indeed, deer mice form the foundation of much of our understanding of the biology of small mammals. The multitude of ecological conditions to which deer mice have adapted has contributed to an impressive array of biological diversity within a single, ubiquitous genus. While this radiation is fascinating in its own right, *Peromyscus* is arguably foremost among nascent model systems that may aptly model the genetic complexity of the human condition, which too has long been shaped by natural selection in the wild. We hope that the continued development—primarily through the growth of genetic and genomic resources—of this model system rooted in ecology will galvanize research in all corners of biology.
1. **Much ecological diversity remains untapped.** The bulk of our understanding of *Peromyscus* biology comes from studies of two ubiquitous species that have proven amenable to laboratory life—*P. maniculatus* and *P. leucopus*. However, the majority of *Peromyscus* species remain comparatively understudied, particularly in Central America and Mexico where taxonomic diversity and endemism is greatest.

2. **The genus is in need of taxonomic revision and genomic resources.** A comprehensive phylogeny based on genome-wide DNA sequences would greatly facilitate the comparative approaches that are the unique advantage of the *Peromyscus* system. An annotated genome assembly is currently available for *P. maniculatus bairdii* (Pman_1.0, GenBank assembly accession GCA_000500345.1) and draft sequences are available for *P. californicus*, *P. leucopus* and *P. polionotus* (Baylor College of Medicine, www.hgsc.bcm.edu/Peromyscus-genome-project), but several more are forthcoming and still more are needed.

3. **Peromyscus may be more applicable to biomedicine than previously thought.** Indeed, certain aspects of human biology—including aging, epigenetics, retinal development and haematology—have been suitably modeled in *Peromyscus* (e.g. Ungvari et al. 2008, Shorter et al. 2012, Sun et al. 2014, Arbogast et al. 2013). Continued research could identify additional areas in which genetically diverse *Peromyscus* may complement biomedical studies of traditional laboratory species.


29


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.

Niko Tinbergen
2.1 Abstract

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, while 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, while *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, pups of both species were reciprocally cross-fostered. Fostering did not alter the characteristic burrowing behavior of either species, suggesting 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 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.

2.2 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 P24). We found striking interspecific differences in both the timing and progression of burrow construction (Figure 1; Table A1). Notably, *P. polionotus* were precocious diggers, constructing complete burrows—defined as excavations with at least two components: an entrance tunnel plus 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 A1 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 A1).

Whereas tunnel length increased with age in both species, reflecting a progression in burrowing ability with growth and development (Figure 1c; 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 x 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 when 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 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. [7]), we examined general measures of morpho-
logical 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 travelled less distance in a 90-minute wheel-running assay than *P. maniculatus*. While 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 A1). Likewise, we did not observe heterochrony favoring *P. polionotus* with respect to additional developmental milestones, as *P. maniculatus* reached them earlier in life (Figure A1). Thus, precocious burrowing in *P. polionotus* juveniles reflects a behavioral difference, 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, pups were reciprocally cross-fostered 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.

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). Following the onset of burrowing, fostered animals constructed burrows no more frequently (4 of 14 mice) than pups reared by their biological 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
Figure 2.2: Reciprocal interspecific cross-fostering. (A) Schematic of cross-fostering design with *P. maniculatus* (yellow), *P. polionotus* (blue) and cross-fostered pups highlighted in red. (B, E) Proportion of mice constructing complete burrows, (C, F) proportion of mice building an escape tunnel, and (D, G) length of excavations. 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, by ANCOVA (see text for details). Significance levels: p ≥ 0.05 = not significant (ns).

(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 following 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 x \( P. \) maniculatus experimental cross (Figure 3a).

The development of burrowing behavior in first generation (F1) hybrids closely matches \( P. \) 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 (F1) vs. 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 (F1) vs. 8 of 22 mice (\( P. \) polionotus)), and the length of excavations (Figure 3c; ANCOVA, \( p = 0.115; \) Cohen’s \( d = 0.68 \)). Moreover, F1 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 (F1) vs. 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 (F1) vs. 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 (F1 hybrid adults build \( P. \) polionotus-like burrows with regard to both length and shape \([2,4]\)).
Figure 2.3: Genetic dissection of precocious burrowing in *P. polionotus* x *P. maniculatus* hybrids. (A) Schematic of breeding design showing *P. maniculatus* (yellow), *P. polionotus* (blue), first-generation F1 hybrids (green) and second-generation backcross (BC) hybrids (grey). (B) Proportion of juvenile animals digging complete burrows (upper panel) and escape tunnels (lower panel); groups compared using Fisher’s exact tests (see text for details). (C) Length of excavations in F1 hybrids compared to *P. maniculatus* and *P. polionotus*; species differences were evaluated by ANCOVA (see text for details). Sample sizes for each group are shown below. (D) Timeline of the four behavioral assays completed for each BC hybrid. (E) Excavation length at juvenile and adult time points. Shading indicates whether each individual was a precocious (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 at ± 1 SE of the mean. 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 means for each group. Significance levels: p ≥ 0.05 = not significant (ns), p ≤ 0.05 = *, p ≤ 0.01 = **, p ≤ 0.001 = ***.
A *P. polionotus* allele affects both juvenile onset and adult expression of burrowing behavior

To test if developmental traits (namely, precocious burrow construction) and adult traits (long entrance tunnels, 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 trials (P61 and P64) (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, while 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 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 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 if 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 A2). We found that inheritance of a *P. polionotus* allele on linkage group 2 was significantly associated with variation in burrowing.
Figure 2.4: 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 at ± 1 SE of the mean. The mean trait values for each parental species are plotted as horizontal bars: *P. polionotus* (blue) and *P. maniculatus* (yellow). 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, B, C) or two adult (A, 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, D) are: p ≤ 0.05 = *, p ≤ 0.01 = **. See also Figure A2 and Table A2.

behavior at both juvenile and adult stages (Figure 4a; linear mixed-effects model, p = 0.02, post hoc Benjamini-Hochberg correction with 10% 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-effects 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 mixed-effects model, p = 0.0303). Moreover, genotype at this marker explains variance in each behavior: precociousness score (6.4% PVE), juvenile excavation length (10.9% PVE), and adult excavation length (8.2% PVE) [8]; 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 A2; \( p > 0.05 \)), 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.

2.3 Discussion

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 to successfully tease 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 [9].

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. [10-13]), or expose underlying proximate mechanisms driving changes in behavior (e.g. [14-17]). 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 (<20cm), simple burrows as adults, but one species, *P. polionotus*, has recently evolved a stereotyped burrowing behavior that results in a considerably longer burrow (>100cm in the wild) comprised 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, 18-21]; Figure 1a). The burrows of *P. polionotus* have inspired studies of phylo-
genetic 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. [22,23]). 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 [24], 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 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 adult burrows, they are smaller in overall size, likely due to the energetic cost of burrowing.

Using a cross-fostering experiment, we next tested if 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. [25]). We found that cross-fostering results do not differ if single or multiple pups are transferred to heterospecific parents, suggesting 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 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—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 [26,27]. On the other hand, because changes in several traits are often involved during adaptation to a new environment [28-30], co-inheritance of groups of phenotypes (e.g. by pleiotropy or linkage) can expedite adaptation [31-33]. Indeed, a common experimental outcome is to map multiple traits to a shared genomic region [34-39], 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. While variation in adult burrows can affect fitness [23,40], juvenile burrowing behavior may also be a target of selection. For example, natural selection for earlier burrowing in *P. polionotus* may reflect (i) its open habitat [18], which may expose young mice to predation and thus increase the survival value of burrowing, or (ii) a form of “play” during a critical period of motor development.
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. [44-47]). 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 [48].

### 2.4 References


3

Behavioural mechanisms of cooperative burrowing in *Peromyscus* mice
3.1 Abstract

While some behaviours are largely fixed and invariant, others can respond flexibly to different social contexts. Here, we leverage the unique burrowing behaviour of deer mice (genus *Peromyscus*) to investigate if and how individuals of three species adapt their behaviour when digging with partners. First, we find that pairs of mice from monogamous, but not promiscuous, species cooperatively construct burrows that are approximately twice as long as those dug by individuals, more similar in size to burrows found in the wild. Then, using a novel behavioural assay in which we can directly observe and measure burrowing behaviour of both individuals and pairs, we find that opposite-sex pairs of monogamous mice construct longer burrows than same-sex pairs. Surprisingly, however, longer burrows are achieved, not by changing individual behaviour, but instead because opposite-sex pairs are more socially cohesive and thus more likely to dig simultaneously – a more efficient mode of burrow elongation. Thus, across social contexts, individual burrowing behaviour appears largely invariant, even when the resultant burrow from pairs of mice differs from the expectation based on individual behaviour, underscoring the fixed nature of burrowing behaviour in *Peromyscus*.

3.2 Introduction

Animal behaviours, like other heritable traits, are subject to evolution by natural selection, yet the conditions under which selection promotes rigid or flexible expression of a given behaviour are not well understood. For some species, the reliable execution of a fixed response can be advantageous,
while, for others, the ability to flexibly adapt to changing contexts may be preferred. Ernst Mayr described these fixed and flexible behaviours as being governed by “closed” (i.e., impervious to modification) versus “open” genetic programs (i.e., subject to external influences). The same behaviour may be “closed” in one species, yet “open” in another. For instance, species recognition in parasitic cowbirds—a species reared by heterospecific foster parents—relies on a fixed genetic template to identify conspecifics, whereas greylag geese, famously, imprint upon the first object they encounter, conspecific or otherwise. Mayr suggested that “open” behavioural programs may be more likely to evolve in social species, in which flexibility in complex social environments may be especially beneficial. However, comparative studies, in which homologous behaviours are contrasted among species with different social systems, are lacking.

Animal architectures – such as termite mounds, bird nests, or rodent burrows – are stereotyped structures that can aid in species identification or the reconstruction of phylogenetic histories and, as such, are prime candidates for comparative studies of homologous behaviour. The burrows built by North American deer mice (genus *Peromyscus*) are strongly heritable and can be readily quantified in the laboratory. For example, the oldfield mouse (*P. polionotus*) constructs “complex” burrows consisting of a long entrance tunnel, central nest chamber, and upward sloping escape tunnel. By contrast, its sister species (*P. maniculatus*) and an outgroup (*P. leucopus*) produce simple burrows comprised of only a short entrance tunnel and terminal nest chamber. Moreover, *P. polionotus* cohabitates in opposite-sex pairs that are both socially and genetically monogamous, whereas *P. maniculatus* and *P. leucopus* nest solitarily and are highly promiscuous. Here, we leverage this natural variation in mating system to explore the relationship between sociality and behavioural plasticity.

Using a combination of behavioural assays – that capture both variation in burrow architecture and digging behaviour – we tested whether mice alter their behaviour when burrowing alone versus with a partner. More specifically, in three *Peromyscus* species, we compared the lengths of burrows

52
constructed by individuals and pairs, as well as compared burrows constructed by same-sex and opposite-sex pairs. We show that mice of only the monogamous *P. polionotus* dig cooperatively, and that opposite-sex pairs dig longer burrows than same-sex pairs. However, we find that opposite-sex pairs do not spend any more time digging than same-sex pairs, but rather that opposite-sex pairs are more socially cohesive and more likely to dig simultaneously, which is a more efficient mode of burrow elongation. Thus, while cooperative burrowing has evolved in only the monogamous species, individual burrowing behaviors, such as time spent digging, appear to be conserved across social contexts in all three species.

3.3 Results

Burrow length varies with social context in a monogamous species

When tested as individuals, *Peromyscus* mice build burrows that are highly stereotyped and species-specific in size and shape (Fig. 1a). However, in the wild, burrows can be longer than those produced in the lab and thus, may be constructed by more than one individual. Whether the burrows constructed by pairs of mice differ in size or shape from individual burrows remains unknown. To address this question, we first compared the length of the longest burrow produced by individuals and pairs of mice. While species-specific burrow shape remained unaltered between individual and pair trials, we found significant variation in burrow size. Using linear mixed-effects models (LMMs) that control for sex and mouse identity, we found that individuals and pairs dug equivalent burrows in *P. leucopus* and *P. maniculatus*, but that pairs of mice dug significantly longer burrows than individuals in *P. polionotus* (Fig. 1b, LMM, *P*. leu: *P* = 0.631; *P*. man: *P* = 0.071; *P*. pol: *P* < 2e-16). In *P. polionotus*, a monogamous species, pairs dug 82% longer burrows on average than individuals (64.3 ± 4.2 cm vs. 35.3 ± 2.1 cm). Thus, the response to social context varies by species: the more so-
cial species (*P. polionotus*) produces longer burrows in pairs, whereas its more solitary congeners (*P. maniculatus* and *P. leucopus*) do not.

Because sex—of both the individual and the social partner—can be an important source of variation in behaviour, we next tested for sex differences in burrow length, both between individual males and females and between same-sex and opposite-sex pairs. In *P. leucopus*, we failed to detect significant burrow length differences between sexes or pair-types (Fig. 1c, LMM, sex: *P* = 0.418, pair-type: *P* = 0.467). By contrast, individual males dug longer burrows than individual females in both *P. maniculatus* and *P. polionotus* (Fig. 1c, LMM, *P*. man: *P* = 0.027; *P*. pol: *P* = 0.001). However, only in *P. polionotus* did we observe variation in burrow length between pair-types, with opposite-sex pairs producing longer burrows than same-sex pairs (Fig. 1c, LMM, *P* < 0.001).
Figure 3.1: *P. polionotus* pairs cooperatively construct burrows that are nearly twice as long as those dug by individuals. 

**a**. Schematic of phylogenetic relationships among three *Peromyscus* species and their characteristic burrow architectures. 

**b**. Ratio of pair to individual burrow lengths in three species. A ratio >1 indicates that two mice together dig a longer burrow than one mouse alone. All pair-types (FF, MM, FM) and both sexes (F, M) are included. In *P. leucopus* and *P. maniculatus*, pairs and individuals dig burrows of the same length (*P. leu*: LMM, $P = 0.631$, $n = 9$ pairs, 15 singles; *P. man*: LMM, $P = 0.071$, $n = 22$ pairs, 21 singles). In *P. polionotus*, pairs dig 82% longer burrows than individuals (LMM, $P < 2\times10^{-16}$, $n = 52$ pairs, 52 singles). 

**c**. Individual and pair burrow lengths across species. Data points represent the mean of up to 7 trials per individual or unique pair (mean 2.5). We detected sex differences in both *P. maniculatus* and *P. polionotus* individuals (light purple vs. light green). We detected pair-type differences only in *P. polionotus* (dark purple vs. dark green vs. gold). 

**d**. Ratio of observed to expected pair burrow lengths, given the known output of individuals comprising the pair. Significance levels: *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$. Error bars represent SEM.
To ascertain whether sex differences in burrowing behaviour can explain the observed variation among *P. polionotus* pair-types, we tested whether burrows dug by different pair-types were shorter or longer than expected, given the individual burrow lengths of animals comprising the pair. In *P. polionotus*, individual males dug on average 36% longer burrows than individual females (40.9 ± 3.1 cm vs. 30.1 ± 2.6 cm) (Fig. 1c). For each *P. polionotus* pair-type, we calculated the observed:expected burrow length ratio, in which the observed value is the mean burrow length dug by a given pair, and the expected value is the sum of the mean burrow lengths produced individually by each member of the pair. We found that opposite-sex pairs dug longer burrows than expected relative to same-sex pairs (Fig. 1d, LMM, *P* = 0.005). The observed:expected ratio was significantly less than 1 for same-sex pairs, but not significantly different from 1 for opposite-sex pairs (Fig. 1d, t-tests, FF: *P* = 0.012; MM: *P* = 0.006; FM: *P* = 0.234), suggesting that per capita output declines with a same-sex partner but is unchanged with an opposite-sex partner. This raises the possibility that mice invest fully when constructing a burrow intended for reproduction, but not when constructing a burrow intended for shelter.

**P. polionotus** females increase burrow length after cohabitation with a male

We next asked if both mice in an opposite-sex pair invest equally, or if one sex predominately contributes to the construction of a shared burrow. Early studies of burrowing in *P. polionotus* hypothesized that reproductive females may be responsible for the majority of burrow construction\(^{10}\). Accordingly, we asked if females increase their burrow output in response to—or in anticipation of—a change in reproductive status, and, if so, can this explain the observation that opposite-sex pairs construct longer burrows?

To determine if females with an opportunity to reproduce increase their burrow output, we used a repeated measures design to compare the burrows dug by individual females before and after being co-housed with a conspecific male (Fig. 2a). We tested both *P. maniculatus* and *P. polionotus*. 

56
Figure 3.2: *P. polionotus* females dig longer burrows after cohabitation with a male, regardless of pregnancy status. Experimental design schematic. Individual females were first tested as virgins to determine baseline burrowing output. After a period of cohabitation with a male—during which approximately half of subjects became pregnant—females were assayed again. a, After cohabitation with a male, 15/19 (79%) of pregnant *P. maniculatus* females produced longer burrows, compared to only 6/17 (35%) of non-pregnant controls. The median burrow length increase among pregnant *P. maniculatus* females was 21% (P = 0.023, n = 19). By contrast, females that did not become pregnant after a period of male cohabitation showed no change, on average, in burrow length relative to their virgin trials (least-squares means, P = 0.787, n = 17). c, Regardless of pregnancy status, 26/42 (62%) of *P. polionotus* females produced longer burrows after cohabitation with a male, and the median increase in burrow length relative to virgin trials was 11% (LMM, P = 0.020, n = 26). Each line represents the mean of 2 trials per individual, per stage (i.e., before and after male cohabitation). Significance levels: *P < 0.05. Error bars represent SEM.

by generating LMMs controlling for female age and length of the cohabitation period. In both *P. maniculatus* and *P. polionotus*, females with longer virgin burrows were no more likely to become pregnant than those with shorter virgin burrows, discounting a relationship between burrow output and fertility (binary logistic regression, *P. man*: P = 0.836; *P. pol*: P = 0.979).

In *P. maniculatus*, pregnant but not control (i.e., non-pregnant) females dug longer burrows relative to their virgin trials (13.6 ± 1.1 cm vs. 10.9 ± 0.9 cm) (Fig. 2b, least-squares means, pregnant: P = 0.023, control: P = 0.787), indicating that pregnancy, and not simply cohabitation, drives an increase in burrow output. By contrast, cohabitation with a male significantly impacted the burrow length of *P. polionotus* females, regardless of pregnancy status. Post-cohabitation females dug slightly longer burrows (26.2 ± 2.6 cm vs. 22.0 ± 1.4 cm), with a median increase of 11% over their virgin trials.
Together, these results indicate that individual females burrowing in a reproductive context experience changes in burrow output. However, in *P. maniculatus* this change is likely mediated by the hormonal shifts associated with pregnancy, whereas in *P. polionotus* the increase in burrow length is associated with previous social experience (i.e., cohabitation with a male). We next asked if this 11% increase in burrow output by females exposed to males could explain the longer-than-expected opposite-sex burrows observed in *P. polionotus* (Fig. 1c). However, even after accounting for this increase, we still find a significantly greater observed:expected burrow length ratio for opposite-sex than same-sex pairs (LMM, *P* = 0.012). This raises the possibility that males, as well as females, increase burrowing effort after cohabitation with an opposite-sex partner, or that burrowing with an opposite-sex partner alters individual digging behaviour in a way that cohabitation alone does not.

**Novel behavioural assay reveals sex-specific responses to burrowing in pairs**

To identify the behavioural mechanisms by which *P. polionotus* mice produce burrows of different lengths, we designed a narrow and transparent sand-filled enclosure in which we could directly observe and measure burrowing behaviour (Fig. 3a). With this assay, we quantified individual digging behaviours and measured real-time progress in burrow construction (see Methods). Using a repeated measures design (Fig. 3b), we asked if *P. polionotus* individuals alter their burrowing behaviour when digging with a same-sex versus opposite-sex partner. Specifically, we asked if the longer burrows produced by opposite-sex pairs result from increased effort (i.e., digging duration) and/or increased efficiency (i.e., burrow extension rate). First, we measured the total burrow length dug by individuals and pairs of mice and tested for differences using LMMs controlling for trial number and mouse identity. For individual females, burrowing with an opposite-sex, but not same-sex, partner significantly increased total burrow length (Fig. 3c; Tukey’s test, *F* vs. *FF*: *P* = 0.219; *F* vs. *FM*: *P* < 1e-4). However, for individual males, burrowing with a partner of either sex signifi-
Figure 3.3: Sex differences in digging behaviour across social contexts. a, Schematic of the behavioural apparatus and snapshot of two mice digging in tandem. b, Experimental design. Baseline burrowing output was first quantified in virgin females and virgin males. Mice were then assayed—in random order—with both a same-sex and an opposite-sex burrowing partner. c-d, Repeated measures plots of burrow length across three social contexts: individual, same-sex, and opposite-sex trials. Each line represents the mean of 2 trials per mouse, per context. c, Two females did not dig significantly longer burrows than a single female, but female-male pairs dug longer burrows than two females (Tukey’s test, P = 0.002). d, Two males dug significantly longer burrows than a single male (Tukey’s test, P < 0.001), but burrows dug by female-male pairs did not differ significantly from those dug by two males. e-g, Data points represent the mean of 8 observations per individual. Benjamini-Hochberg adjusted p-values are reported. e, Total time spent underground and total time spent digging, per 10-minute observation period. f, Number of underground and digging bouts, per observation. g, Mean duration of underground and digging bouts, per observation. h-j, Repeated measures plots of individual-level burrowing behaviours during same-sex and opposite-sex trials. Each line represents the mean of 4 observations per individual, per trial-type. h, Total digging duration. i, Mean underground bout duration. j, Mean digging bout duration. Significance levels: *P < 0.05, **P < 0.01, ***P < 0.001. Error bars represent SEM.
icantly increased total burrow length (Fig. 3d; Tukey’s test, M vs. MM: P < 0.001; M vs. FM: P = 0.003). These results suggest that, while males may burrow cooperatively with a partner of either sex, females may be more likely to do so with an opposite-sex partner.

Next, to determine a baseline for individual-level behaviour, and to assess overall sex differences, we pooled data across same-sex and opposite-sex trial-types. Controlling for social context, males spent more time underground (Fig. 3e, LMM, P < 0.001) and more time digging than females (LMM, P < 0.001). Males also entered the burrow more often (Fig. 3f, LMM, P = 0.021) and initiated more digging bouts (i.e., defined digging sessions) than females (LMM, P < 0.001). Overall, the mean duration of an underground bout did not differ between males and females (Fig. 3g, LMM, P = 0.053), but males had significantly longer digging bouts (LMM, P = 0.002). Together, these behavioural differences are consistent with our previous observation that individual males dig longer burrows than individual females (Fig. 1c). In summary, males produced longer burrows both by engaging in digging sessions more frequently and by digging for longer once those sessions were initiated.

Whereas the total time spent digging was split evenly between each member of a same-sex pair, the division of labour was significantly skewed in opposite-sex pairs (Fig. S1a, LMM, P = 1.68e-7). On average, females performed 32% of the total pair digging, while males contributed the remaining 68%. This imbalance can be entirely explained by the disparity in time spent digging between individual males and females, which is unaffected by social context (see below).

Because males spend more time digging than females, we might expect—under a simple additive model—that two males would spend the most time digging and produce the longest burrows. However, consistent with our burrow length comparisons (Fig. 1c), opposite-sex pairs spend just as much time digging (Fig. S1b) and produce burrows that are just as long (Fig. S1c) as two males.

Social cohesion, but not individual digging behaviour, changes with context
Thus, to determine if social context affects individual behaviour, we next asked if males or females altered their digging behaviour between same-sex and opposite-sex trials. Notably, we did not detect any differences in the number of digging bouts (LMM, P = 0.112) or total digging duration between trial-types for either sex (Fig. 3h, LMM, P = 0.340). Intriguingly, mean underground bout duration was the only metric for which we detected significant differences between social contexts: on average, females had 53% longer underground bouts in opposite-sex than same-sex trials, while males remained unaffected (Fig. 3i, least-squares means, females: P = 0.002, males: P = 0.999). Additionally, both sexes had significantly longer digging bouts—10% on average—in opposite-sex versus same-sex trials (Fig. 3j, LMM, P = 0.039). Together, these results suggest that females spend more time in the burrow when a male rather than a female is present, and that mice of both sexes respond to an opposite-sex burrowing partner with slightly longer digging bouts.

To identify a behavioural mechanism responsible for differences in burrow length among pair-types, we tested for differences in social behaviour that might indicate greater cohesion in opposite-sex pairs. Indeed, opposite-sex pairs engaged in more affiliative behaviours (e.g., allogrooming, huddling) than same-sex pairs (Fig. 4a, LMM, P = 0.004). Furthermore, we observed agonistic interactions (e.g., boxing, biting, fleeing) in 44% of two-male trials, but only in 28% of opposite-sex trials. Both were significantly greater than two-female trials, in which no agonistic interactions were observed (Fig. 4b, Fisher’s exact test, MM vs. FF: P = 0.0073; FM vs. FF: P = 0.041). These results raise the possibility that differences in social cohesion among pair-types may reflect differences in the propensity to burrow cooperatively.

**Opposite-sex pairs are more likely to engage in efficient simultaneous digging**

To understand how opposite-sex pairs generated longer burrows despite no significant increase in total digging duration (Fig. 3h), we tested for differences in efficiency among pair-types. To calculate digging efficiency (i.e., burrow extension rate), we divided the change in burrow length over the
course of each 10-minute observation period by the total digging duration in that period (see Methods). We distinguished between independent digging (i.e., mice working separately, either non-concurrently or on different burrows) and simultaneous digging (i.e., both mice working on the same burrow at the same time) (Video S1). As expected, the duration of both independent digging and simultaneous digging were significant predictors of change in burrow length. We calculated partial correlation coefficients both for independent and simultaneous digging (Fig. 5a, b, independent digging: $r = 0.255$, $P < 0.001$; simultaneous digging: $r = 0.356$, $P = 8.14e-7$). To control for the number of mice, independent and simultaneous digging were both expressed in man-hours or, more accurately, mouse-minutes (i.e., time spent digging multiplied by the number of mice digging). We found that an additional mouse-minute of independent digging resulted in an additional $0.46 \pm 0.13$ cm of burrow length. However, an additional mouse-minute of simultaneous digging resulted in an additional $1.10 \pm 0.21$ cm, indicating that concurrent, coordinated digging is a more efficient mode of burrow extension, even when controlling for total digging duration (Fig. 5c, multiple linear

Figure 3.4: Opposite-sex pairs have more affiliative and fewer agonistic interactions. a, Total number of affiliative interactions observed per unique pair. Each data point represents the sum of 4 observations per pair. b, Percentage of trials in which agonistic behaviour between partners was observed. Significance levels: *$P < 0.05$, **$P < 0.01$. Error bars represent SEM.
We then asked if the probability of simultaneous digging differed among pair-types. We found that opposite-sex pairs were more likely to engage in simultaneous digging, even after controlling for time spent digging by both the focal mouse and its partner (Fig. 5e, binary logistic regression, \( P = 0.025 \)). The probability of observing simultaneous digging was significantly higher for opposite-sex pairs.
sex than same-sex pairs: 54 ± 5% for opposite-sex pairs, but only 39 ± 7% for two males and 32 ± 8% for two females (Fig. 5f, binary logistic regression, P = 0.025). Together, our results suggest that longer opposite-sex burrows can be explained predominately by an increase in burrowing efficiency (mediated by enhanced social cohesion) rather than an increase in individual burrowing effort (i.e., digging duration).

3.4 Discussion

Our study demonstrates first that pairs of mice from the monogamous species build significantly longer burrows than individuals that burrow length is modulated by social context in a monogamous species with biparental care, but not in two closely-related promiscuous species with uniparental care. These results are consistent with the prediction that natural selection should favour “open” behaviour programs in species with greater sociality, where there is more opportunity for learning and integrating experiential information. In the wild, P. polionotus is an obligate and often social burrower, whereas both P. leucopus and P. maniculatus are largely facultative, solitary burrowers. These differences in ecology, mating system, and sociality likely contribute to different selective pressures shaping the behaviour program of each species.

For example, both P. leucopus and P. maniculatus are known to nest in a variety of locations, including rock crevices, brush piles, fallen logs, and tree cavities. This diversity of potential refuges hints that these mice may only infrequently excavate a burrow de novo. Moreover, radiotelemetry and nest-box studies indicate a general pattern of solitary nesting in both species, particularly during the breeding season. Because alternative nest sites are often available, selection may favour a goal-oriented strategy in which mice dig only until a minimum habitable space is achieved. This interpretation is supported by our observation that burrows built by P. leucopus pairs were no longer than those dug by individuals, and that P. maniculatus pairs—of all types—produced burrows that
were only slightly longer than those built by individual conspecifics (Fig. 1b).

By contrast, a different strategy may govern the expression of burrowing behaviour in *P. polionotus*, a species that nests almost exclusively in burrows of its own construction. We found that *P. polionotus* pairs produced burrows that were nearly twice as long as those built by individuals (Fig. 1b), suggesting that this species employs a more effort-oriented strategy in which mice dig until some amount of individual effort has been expended. This interpretation also supports our finding that social modulation of burrow length in *P. polionotus* arises largely independently of changes to individual digging duration (Fig. 3h).

A key finding of our study is the observed per capita increase in burrow output with an opposite-sex partner in *P. polionotus*. Dawson et al.\textsuperscript{10} assayed mated pairs with nursing litters and hypothesized—based on field observations and preliminary laboratory trials comparing males with both virgin and lactating females—that the female was responsible for the majority of burrow construction. We therefore predicted that females, primarily, would upregulate their burrow output in response to—or in anticipation of—a change in reproductive status. However, we found only marginal increases in burrow output after male cohabitation and no significant differences between pregnant and non-pregnant *P. polionotus* mice. By contrast, *P. maniculatus* females dug significantly longer burrows while pregnant, suggesting that selection in this species may favour strong behavioural drive to burrow independently during pregnancy. Indeed, field studies indicate that *P. maniculatus* females frequently relocate their nest sites at parturition and during early lactation, ostensibly as a predator-avoidance tactic\textsuperscript{22,26,27}. Conversely, the marginal increases in *P. polionotus* burrow output appeared to be modulated by social experience rather than hormonal shifts associated with pregnancy. Pregnant *P. polionotus* may not upregulate to the same extent as *P. maniculatus* because burrow construction is typically shared between both members of a monogamous pair\textsuperscript{16,28}.

Collective building behaviour has evolved in diverse taxa, from social spiders\textsuperscript{29} and eusocial insects\textsuperscript{30}, to communally nesting birds\textsuperscript{31} and cooperatively burrowing rodents\textsuperscript{32}. Two levels of
causation—ultimate and proximate—have been offered in explanation of these complex "extended phenotypes". Kin selection\textsuperscript{33} is commonly invoked as an ultimate cause of cooperation in animal societies, yet the mechanisms that maintain cooperation among non-kin are still debated\textsuperscript{34}. One proposed explanation is that seemingly cooperative behaviours arise as “by-product mutualisms” in which animals act to maximize their own immediate fitness and, in the process, impart benefits to other individuals\textsuperscript{35}. Here, we demonstrate that unrelated and unfamiliar pairs of mice readily construct shared burrows that, in the wild, would have direct fitness benefits\textsuperscript{36}. Concurrent with the mutualism hypothesis, even unrelated same-sex pairs in our study engaged in some amount of socially coordinated digging, suggesting that the prospect of future reproduction is not necessary for cooperative behaviour.

Proximate explanations, meanwhile, address how complex architectures emerge from individual-level actions, and, relatedly, how such structures adapt to changing environments or social contexts. For most ants and termites, nest volume correlates strongly with colony size\textsuperscript{37}, and such adaptive scaling can be accomplished through a negative-feedback process in which insects adjust their digging rate according to traffic flow, or the density of workers in the nest\textsuperscript{38}. Similarly, the topography of active building sites within the nest is regulated by “cement pheromone”—a stigmergic cue that induces soil deposition by workers\textsuperscript{39,40}. The pheromone evaporates more quickly under hot, dry conditions and causes a change in the spacing of nest pillars, likely protecting the colony from desiccation\textsuperscript{42}. These examples from social insects illustrate that adaptive plasticity in nest architecture can be achieved through the implementation of simple rules of construction, without direct communication among individuals\textsuperscript{43}. Here, we argue that \textit{P. polionotus} mice likely achieve flexibility in burrow size through similarly simple mechanisms, such as digging in concert with an opposite-sex, but not same-sex, partner.

Our initial prediction—that individuals actively increase their digging duration in response to an opposite-sex digging partner—was not supported (Fig. 3h). Rather, elongated burrows emerge
by the simple means of greater spatiotemporal overlap between socially cohesive opposite-sex pairs. We showed that opposite-sex pairs are more likely to dig concurrently in the same burrow—a more efficient mode of burrow elongation that produces overall longer burrows despite no change in total digging duration (Fig. 5e, f). This behaviour is reminiscent of “chain digging” observed in other communal burrowers such as eusocial mole rats and group-living degus and suggests that distantly related rodent species have converged upon a common strategy for the efficient excavation of shared living space.

Conclusion

Here, we report that P. polionotus—a relatively social Peromyscus species—does indeed display more plasticity in burrow architecture, with burrow length varying according to social context. This would suggest a putatively “open” behaviour program, in which animals flexibly adapt their burrow output to suit the current environment. However, we demonstrate that flexibility in the extended phenotype arises not through active modulation of individual digging duration, but rather through a relatively passive mechanism in which socially cohesive opposite-sex pairs engage in more efficient, simultaneous burrowing. Our study highlights the importance of dissecting the individual-level behaviours that collectively produce variation in communal structures.

3.5 References


The adaptive value of burrowing behavior in the wild
4.1 Abstract

Complementary to any study of behavioral genetics in the lab is careful quantification of behavior in the wild. In nature, *Peromyscus* burrowing is a nuanced social behavior and full expression of the underlying genetic modules may occur reliably only under fully natural conditions. It is therefore essential to understand both the social and ecological contexts of burrowing behavior if we are to correctly interpret its function. Here, we employ a novel RFID-tracking system to non-invasively monitor the nightly activity patterns of a population of Santa Rosa beach mice (*P. polionotus leuccephalus*). We installed antennas at burrow entrances and recorded each instance of an RFID-tagged mouse entering or exiting a burrow over a period of 11 nights. From these data, we describe spatial and temporal patterns of burrow usage for individual mice and infer patterns of spatiotemporal overlap for groups of mice. First, we show that groups of mice consistently visit the same network of spatially-clustered burrows. Second, we show that mice overlap at burrow entrances more often than expected by chance, indicating that individuals in the population associate non-randomly. Third, we show that close relatives are more likely to visit the same burrow sites during the night—and to share the same burrow during the day—than more distantly related individuals. Thus, sharing a network of burrows with related individuals in the population may provide advantages in terms of inclusive fitness. Last, we measure the interior temperature of burrows of varying lengths and demonstrate that longer burrows provide better buffering against external temperature fluctuations. Together, we demonstrate that, in addition to providing a retreat from predators and severe climatic conditions, burrows are arenas for social dynamics and may confer benefits to close relatives. Thus, while controlled laboratory assays are important for understanding behavior, observations of
natural populations can shed light on the selective pressures that shape behavioral diversity.

4.2 Introduction

Natural variation in animal behavior has important consequences for fitness in the wild. However, there are few examples of behaviors for which we understand both the genetic basis and the adaptive value. One emerging exception is burrowing behavior in deer mice (genus *Peromyscus*). Deer mouse species are broadly distributed across North America and display a diversity of burrow architectures. Most *Peromyscus* species that have been tested construct short, simple burrows, whereas *P. polionotus* uniquely constructs a complex burrow consisting of a long entrance tunnel, nest chamber, and escape tunnel. These differences in burrow architecture are heritable and have a relatively simple genetic basis. The different environments these animals occupy may offer some clues as to the adaptive significance of their distinctive burrowing behaviors. Species that construct short burrows—such as *P. leucopus* and *P. maniculatus*—live in structured environments such as grasslands or forests, whereas *P. polionotus* lives in open environments such as fallow fields. Thus, *P. polionotus* burrows, which are evolutionarily derived, are likely to be essential for providing protection from predators and severe climatic conditions.

This is particularly true for the *P. polionotus* subspecies that occupy Florida’s highly-exposed barrier islands. These eight coastal subspecies—collectively known as beach mice—are classified on the IUCN red list as near threatened, endangered, critically endangered, or extinct. Improved knowledge of their behavior in the wild is of interest not only to behavioral geneticists, but also to conservation biologists charged with ensuring the continued viability of these endemic animals. The dearth of other rodent species on barrier islands suggests that interspecific competition for food and home sites is low and that predation is likely to be the most important factor affecting the survival of beach mice. Adaptations to predation pressure may include both morphological and behavioral
traits. Indeed, pale coat color effectively camouflages beach mice inhabiting Florida’s coastal dune habitats. Both the genetics and adaptive value of beach mouse pigmentation have been studied extensively. However, our understanding of behavioral adaptation to predation pressure remains comparatively poor.

The first step toward generating adaptive hypotheses for the evolution of complex burrow architecture requires a detailed understanding of how burrows are used in the wild. Early naturalists offered general descriptions of the size, shape, and location of P. polionotus burrows and inferred burrow occupancy by digging up burrows and capturing their occupants. However, due to the difficulty of describing the subterranean behavior of a nocturnal species, these early studies failed to quantify patterns of burrow usage over space or time. Here, we employ a non-invasive, RFID-based tracking system to record the spatial and temporal patterns of burrow usage in a wild population of Santa Rosa Island beach mice (P. p. leucocephalus) over several nights. Additionally, we measure several attributes of burrow sites to understand if abiotic selection pressures might also contribute to the evolution and maintenance of long, complex burrows.

4.3 Results

To better understand spatial and temporal patterns of burrow usage in nature, we conducted a two-part study in the frontal dunes of Santa Rosa Island, FL (Fig. 1a). First, to characterize spatial relationships among animals—and to obtain a population of RFID-tagged mice for monitoring—we performed a spatially explicit capture-recapture experiment (Fig. 1b). We subsequently monitored the activity of the RFID-tagged population at several burrows over 11 nights (Fig. 1c). Both phases of the study were conducted on a 55,200 m² grid on the southern end of Santa Rosa Island, a barrier island in the Gulf of Mexico (Fig. 1d). Grid locations for multiply-trapped beach mice and RFID-monitored burrow sites are shown (Fig. 1e).
Figure 4.1: Home-range characterization and burrow use patterns as determined by capture-recapture and RFID-monitoring.  

a, Study site image depicting frontal dune habitat. b, Beach mouse (*P. polionotus leucocephalus*) emerging from a trap during the release stage of the capture-recapture experiment. c, Beach mouse emerging from a burrow entrance installed with a custom-built RFID antenna. d, Satellite image depicting a section of Santa Rosa Island, FL (30°23'42" N, 86°43'27" W). The dotted line denotes the trap grid boundary (expanded in e). e, Trap grid captures and RFID-monitored burrow locations. Trap area polygons are drawn for 16/32 individuals that were captured at three or more unique trap locations (n = 7 females, n = 9 males). Grey circles denote RFID burrow locations (n = 40). Circle size is log2 scaled by the total number of RFID pings recorded over 11 nights. f, Half-normal detection function describing the capture probability as a function of distance from the home-range centre. Grey dashed line represents the standard error of the model estimate. Red dashed line represents the scale parameter value ($\sigma$). $g_0$ = intercept, $\sigma$ = scale parameter, $hr$ = 95% home-range radius. Data from November 2017.

**Population density and home-range size**

Using a spatially explicit capture-recapture (SECR) model, we estimated a population density of $254 \pm 48$ mice/km$^2$. The probability of trapping a mouse at its home-range centre was $0.019 \pm 0.003$ and the mean home-range size estimate for the population was $14,212$ m$^2$ (Fig. 1f). Our SECR model estimates are very similar to those reported by Blair$^5$, who trapped *P. p. leucocephalus* between
November 1941 and June 1942. Blair calculated a population density of 253 mice/km² and estimated a mean home-range size of 15,641 m² for beach mice occupying frontal dune habitat on Santa Rosa Island. We also independently estimated utilization distributions for individual mice using RFID relocations and found an average home-range size of 6292 ± 1362 m² (Fig. C1). These estimates are closer to those reported by Swilling and Wooten, who found mean home-range sizes of 6798 ± 703 m² for dispersers and 1934 ± 580 m² for philopatric Alabama beach mice (P. p. ammobates).

We next used our SECR density estimate to determine how well we sampled the population within our 690 x 80m trap grid. Assuming an 80m habitat buffer on each side, our sampled area measured 850 x 240m or 0.204 km² and, given an estimated population density of 254 mice/km², we would expect the population size for our study site to be 52 mice. We therefore sampled approximately 75% of the population (i.e., 39/52 mice).

**Burrow visitation profiles**

We next asked if mice visit multiple burrows and, if so, do consistent groups of mice frequent the same burrows. Accordingly, we plotted as pie graphs the portion of each burrow’s total RFID activity comprised by different mice (Fig. 2a). Indeed, consistent groups of mice frequented burrow networks ranging in size from 4-10 burrows (Fig. 2a). We defined burrow networks by clustering burrows according to their visitation profiles (see Methods). We also found that burrows belonging to the same network are spatially clustered (Fig. 2b). This indicates that mice use multiple burrows within their home-range, and that groups of mice consistently visit the same network of burrows.

**Total activity summaries.** We found no significant difference in the total number of burrows visited by males and females (LMM, P = 0.541). Mice visited between 1 and 11 burrows over the course of the study (mean = 5.7 ± 0.5 s.e. burrows per mouse, n = 26 mice) (Fig. 2c) and we detected a total of 1 to 10 mice at each burrow (mean = 4.1 ± 0.4 s.e. mice per burrow, n = 36 burrows) (Fig. 2d).
Nightly activity summaries. We found no significant difference in nightly activity level (i.e., number of RFID pings) between males and females (GLMM, \(P = 0.631\)). We also found no significant sex difference in the number of burrows visited per night (GLMM, \(P = 0.162\)). Mice visited between 1 and 11 burrows per night (mean = \(2.8 \pm 0.3\) s.e. burrows per mouse, \(n = 26\) mice) (Fig. 3a) and, each night, we detected between 1 and 8 mice at each burrow (mean = \(2.4 \pm 0.2\) s.e. mice per burrow, \(n = 36\) burrows) (Fig. 3b). Together, the total and nightly activity summaries suggest that mice may visit the burrows constructed by their neighbours on a regular basis.

Burrow preferences. Over the course of the study, the majority of mice (24/26) were detected at more than one burrow. These burrows tended to be spatially clustered, but occasionally mice were pinged at burrows nearly 600 m apart. The mean distance between burrow sites in an individual’s burrow network was \(79.7 \pm 13.5\) m (\(n = 24\) mice) (Fig. 3c). However, not all burrows were used equally. On average, 56 ± 5% of a mouse’s total RFID activity was observed at a single
Figure 4.3: Nightly activity summaries, distance between burrows, and burrow use evenness. a, Mean number of burrows visited per night for each mouse (n = 26 mice). b, Mean number of mice detected per night for each burrow (n = 36 burrows). c, Mean distance between all burrows each mouse’s network (n = 26 mice). d, Portion of each mouse’s total RFID activity observed at each burrow in its network (n = 26 mice). Data from November 2017.

burrow (n = 26 mice) (Fig. 3d). Thus, while mice consistently visited several burrows within their home-range—traveling up to 597 m between burrow sites—the majority of their activity was often observed at a single burrow.

Figure 4.4: Distribution of RFID timestamps and fraction of spatiotemporal overlap in the population as a function of time-bin size. a, Histogram of RFID timestamps for female and male mice over the 11-night study. At the midpoint of the study, sunset and sunrise occurred at 4:46 PM and 6:20 AM, respectively. b, Network load (fraction of all possible pairs of mice that are detected at the same burrow entrance within a given time window). The green dot at 1.5 hours denotes the elbow of the curve and, accordingly, the time-bin size used to calculate overlap index for all downstream analyses. Data from November 2017.
Spatiotemporal overlap

RFID pings are continuously distributed between dusk and dawn (Fig. 4a), suggesting that mice actively use burrows throughout the night. We next asked to what extent mice use the same burrows within the same time window. To answer this question, we developed an overlap index (OI) to describe the fraction of a mouse’s total RFID activity that occurs at the same time and place as another mouse (see Methods). We used a 1.5-hour time-bin for all OI calculations (Fig. 4b).

We found evidence of spatiotemporal overlap for 44% (144/325) of possible pairs. The maximum OI was 0.03 and the mean OI (excluding zeroes) was 0.007 (Fig. 5a). To test if mice have greater temporal overlap at burrow entrances than expected by chance, we randomly assigned shuffled...
mouse and burrow IDs to each RFID timestamp before re-calculating OIs. We found evidence of spatiotemporal overlap for 100% (325/325) of pairs in the randomized dataset. However, the magnitude of overlap was much lower than observed in the real dataset. The maximum OI was 0.007 and the mean OI (excluding zeroes) was 0.004 (Fig. 5b). This suggests that mice overlap less often than expected from a random model of mouse movement, but that when they do overlap, pairs coincide to a greater extent than expected by chance. Therefore, mice do not through the landscape independently, but tend to enter and exit burrows within the same time window as other mice.

Social network analysis and genetic relatedness

To better understand social patterns of burrow usage in beach mice, we next used social network analysis. We first constructed a social network using 1.5-hour OIs (Fig. 6a). We then calculated the degree centrality of each node in the network to determine the number of partner mice with which each focal mouse interacted. Mice overlapped with 0-10 other mice in the RFID-tracked population (mean = 5.5 ± 0.6 s.e.). We also found extensive genetic relatedness within the population (Fig. 6b) and that the magnitude of the relatedness coefficient was positively associated with OI (Linear regression, $R^2 = 0.26$, $P = 3.47 \times 10^{-5}$) (Fig. 6c). This suggests that, while mice may be related to several individuals in the population, they tend to share burrows with close kin.

Home burrow analysis. For a subset of mice ($n = 12$), we were able to further determine in which burrow the mouse slept during the day (see Methods). When sleeping location could be confidently assigned, we found that mice consistently spent the day in the same burrow (Fig. 7a), suggesting that these animals have a preferred home burrow. Further, mice that shared a home burrow tended to have higher coefficients of relatedness than those that did not (Linear model, $F = 25.51$, $P = 5.67 \times 10^{-6}$) (Fig. 7b). Mice that shared a burrow had a mean relatedness of 0.40 ± 0.08, whereas mice that slept in different burrows had a mean relatedness of 0.14 ± 0.02 (Fig. 7b). Two cohabitating pairs of adults (F97/M108 in burrow B22 and M104/F105 in burrow F62) had coefficients of
Figure 4.6: Genetic relatedness predicts spatiotemporal overlap between pairs. a, Overlap index (OI) network for the RFID-tracked population. Network nodes represent individual mice, ordered by hierarchical clustering of OI values. Females are indicated in red (n = 11) and males in blue (n = 14). One mouse of unknown sex is indicated in grey. Network edges represent non-zero OIs and are weighted according to OI value. b, Genetic relatedness network. Network nodes represent individual mice are ordered as in a. Network edges are weighted according to the coefficient of relatedness. c, Relationship between genetic relatedness and spatiotemporal overlap. Regression line and standard error are shown in green with grey transparency. Data from November 2017.

Figure 4.7: Close relatives sleep in the same burrow. a, Daytime burrow occupancy for a subset of 12 mice. Colored rectangles denote days in which mice slept in their home burrow. White rectangles denote days in which sleeping location could not be confidently assigned. Mean relatedness coefficients for all pairs within a given burrow are shown at right. c, Relatedness coefficients for pairs of mice that slept in different burrows (white, n = 56) versus the same burrow (grey, n = 10). Significance levels: ***P < 0.001. Error bars represent SEM. Data from November 2017.

relatedness = zero (Fig. 7a). We speculate that these unrelated mice are monogamous pairs.

Burrow site attributes and temperature profiles

The burrows monitored during our RFID study displayed a wide range of physical attributes. Entrance tunnel length ranged from 13-94 cm (Fig. 8a). Burrows were built in a variety of locations; on dune slopes ranging from 9-44 degrees (Fig. 8b) and at elevations ranging from 0-50 feet above sea-
Figure 4.8: Distributions of burrow site attributes for 36 RFID-monitored burrows. a, Entrance tunnel length ranged from 13-94 cm (median = 30.8 cm). b, Burrows were built on dune surfaces with slopes ranging from 9.3-43.5 degrees (median = 21 degrees). c, Burrow site elevations ranged from 0-50 feet above sea level (median = 10 feet). d, Burrow entrances faced nearly all points on the compass, ranging from 40-358 degrees (median = 178). e, Burrow entrances differed in the extent of vegetative cover, with percent cover measurements ranging from 3-99% (median = 58%). f, The only burrow site attribute that correlated with RFID activity was percent cover, with more activity occurring at burrow entrances with more vegetative cover.

level (Fig. 8c). Additionally, burrow entrances faced all directions (Fig. 8d) and varied in the extent of vegetative cover (Fig. 8e). Notably, the extent of vegetative cover at the burrow entrance was a significant predictor of the amount of RFID activity (Linear regression, $R^2 = 0.142$, $P = 0.0196$) (Fig. 8f). Finally, we found that longer burrows are more efficient at maintaining steady temperature profiles. Among artificially constructed burrows, the shortest (10 cm) showed prominent temperature fluctuations, whereas longer burrows (40, 60, 80 cm) showed pronounced climate buffering capabilities (Fig. 9a). We found similar trends when we recorded from abandoned natural burrows (Fig. 9b). These findings point to a thermoregulatory advantage of increased tunnel length.
4.4 Discussion

Essential to any behavioral genetics study is an understanding of the social and ecological contexts of the behaviors we so carefully dissect under controlled laboratory conditions. By pursuing the genetic basis of an adaptive trait, we hope to learn something of its evolutionary history and the selective pressures shaping it. However, until we understand precisely how that trait is expressed in the wild, our analysis is incomplete. Accordingly, the primary objective of this study is to quantify in unprecedented detail how, exactly, burrows are used by individuals and groups of wild mice.

Prior to this study, several basic questions remained unanswered. Do mice use a single burrow or visit a network of burrows? If mice use multiple burrows, are those burrow sites scattered or
spatially clustered? What are the circadian patterns of burrow usage? Do mice overlap at burrow entrances, or do they move independently of one another? Is overlap observed among kin or unrelated individuals? What selective pressures contribute to the extreme burrow length seen in *P. polionotus*?

Taken together, our findings indicate that patterns of burrow usage are more complex than previously thought. Although they do show preferences for particular burrows, beach mice are frequently detected at multiple, spatially-clustered burrows within their home-range. Stable social groups consistently visit burrows within the same network, suggesting that animals may use the burrows constructed by their neighbours as temporary refugia. Further, we did not observe bursts of burrow activity at dusk or dawn, indicating that burrows are actively used throughout the night. We also found that mice show some amount of spatiotemporal overlap with multiple individuals in the population. Interestingly, the greatest overlap is observed among close kin, who also tend to sleep in the same burrow during the day. These findings suggest that the selective advantage of constructing a complex burrow may be conferred to close kin as well as to the builder. In the wild, *P. polionotus* forms monogamous pairs that are stable over time\(^8\), and burrow construction is often shared between both members of the monogamous pair\(^{14,19}\). Thus, while burrow construction is energetically costly\(^20\), the fitness advantages (e.g., environmental buffering\(^{21}\), predator evasion, etc.) may be considerable and shared among relatives.

We also found a strong effect of burrow length on temperature buffering, suggesting that physiological pressures may also play a role in the evolution of complex burrows. Laboratory studies of *P. polionotus* have shown that when physiologically stressed, mice that were permitted to dig burrows had greater survival than those restricted to conventional laboratory cages\(^{21}\). Burrowing mice consumed less water and were better able to maintain body weight under stressful conditions, suggesting that elevated humidity in burrows provides distinct physiological advantages in the form of water economy. Such advantages are compounded when mice live communally. Australian hopping mice experienced 25% less pulmocutaneous water loss when housed in groups rather than singly\(^{22}\).
While, these phenomena have never been explicitly tested in wild *P. polionotus*, our field observations point to a thermoregulatory and humidity buffering function of long burrows.

Ultimately, quantifying selective pressures in the wild will give us a more complete understanding of how behavioral diversity evolves in mammals. This natural history study compliments ongoing laboratory experiments aimed at dissecting burrowing behavior on a genetic and molecular level.

4.5 References


The genetic architecture of burrowing behavior in *Peromyscus* mice: advances from automated phenotyping
5.1 Abstract

Understanding the genetic basis of behavioral diversity remains a major challenge in evolutionary biology. Here, we capitalize on extreme variation in innate burrowing between two sister-species of deer mice (genus Peromyscus) to investigate the behavioral mechanisms and genetic architecture underlying evolved differences in behavior. The ancestral burrow form, exemplified by *P. maniculatus*, is short and simple. By contrast, the evolutionarily-derived burrow form of *P. polionotus* is long and complex. However, we still do not know the precise changes in behavior that give rise to overall differences in species-specific burrow architecture. Here, we use a novel behavioral assay—coupled with automated video analysis and detailed photo quantification—to uncover the behavioral mechanisms underlying variation in burrow architecture. Despite no difference in overall activity, we document that *P. polionotus* spends more of its active time underground. We also find that *P. polionotus* achieves a longer burrow by digging for twice as long, and by digging 5 times more efficiently, than *P. maniculatus*. Furthermore, we show that these component traits map to different parts of the genome, with variation in digging behaviors mapping to chromosome 1 and X, and variation in other movement-related traits mapping to chromosomes 2, 3, and 21. Together, these findings demonstrate that complex behavior can be broken down into its constituent parts to reveal distinct genetic underpinnings for separate components. This provides a useful framework for understanding the role of candidate genes in shaping different parts of the behavior. For instance, natural selection may have targeted genes expressed in different areas of the nervous system, each contributing to different components of burrowing behavior, such as timing or efficiency of digging. Careful quantification of the precise behavioral changes that lead to overall differences in burrow architecture is therefore important, both for the identification and interpretation of candidate genes.
Behavior can be defined simply: it is the repertoire of actions that an organism employs to interact with its environment. However, this straightforward description belies the difficulties inherent in identifying the genetic basis of ecologically-relevant behavior. The challenge is twofold: ecologically-relevant behavior can be difficult to quantify in a controlled fashion, and organisms with well-described behavioral ecology often lack genetic and genomic tools.

First, relative to morphological traits, behaviors can be particularly contingent upon ontogenetic and environmental conditions and, as a result, less reliably expressed. Thus, a significant challenge to the genetic analysis of natural behavior has been accurate quantification, even in a controlled laboratory setting. Generally, the most tractable behaviors are innate (i.e., learning does not play a necessary role), as these are minimally impacted by ontogenetic or environmental contingencies and are likely to have a strong heritable component (Skinner 1966). The structures built by animals (e.g., termite mounds, bird nests, rodent burrows, etc.) are promising candidates for such studies, primarily because these “extended phenotypes” can be precisely measured and because, as the extended products of the actions of genes, they are subject to natural selection (Dawkins 1982).

Second, sufficient genomic resources for the unbiased identification of candidate genes have been lacking in non-model species, thus limiting our ability to attribute behavioral variation to specific genetic variants. As such, behavioral genetics studies have historically been restricted to a handful of genetic model organisms that display behaviors of little or ambiguous ecological relevance (Fitzpatrick et al. 2005). To address these challenges, we 1) develop a novel phenotyping platform to accurately and reliably measure an innate behavior, and 2) employ quantitative trait locus (QTL) mapping to associate this behavior with genetic variation between two non-model species, both with rapidly improving genomic resources.

Our ultimate objective is to identify and understand the role of specific genes in generating natu-
rally evolved behavioral diversity. To this end, we use a forward-genetics approach, to directly identify the chromosomal regions underlying interspecific variation in different components of burrowing behavior between the deer mouse (*Peromyscus maniculatus*) and oldfield mouse (*P. polionotus*). In both species, burrowing behavior is largely innate, critical for survival, and reliably expressed in a controlled laboratory setting. *P. maniculatus* dig short and simple burrows consisting of a single tunnel (Dawson et al. 1988), whereas *P. polionotus* dig stereotyped, long and complex burrows consisting of an entrance tunnel, nest chamber, and escape tunnel (Sumner and Karol 1929, Hayne 1936) (Fig. 1a). Strikingly, mice kept and bred in the laboratory for several generations recapitulate the species-specific burrow architectures observed in nature (Dawson et al. 1988, Weber and Hoekstra 2009, Weber et al. 2013). Furthermore, the complex burrows of *P. polionotus* are derived (Weber and Hoekstra 2009) and likely evolved through changes at a handful of genetic loci, each affecting distinct behavioral modules (i.e., entrance tunnel length and escape tunnel presence) (Weber et al. 2013).

Thus, while previous work has laid a firm foundation, several questions regarding the exact differences in mouse behavior that give rise to overall differences in burrow architecture remain unanswered. For example, does *P. polionotus* excavate longer burrows because it is generally more active? Does *P. polionotus* spend more time digging, or does it dig more efficiently? Quantifying precisely how *P. polionotus* achieves a longer burrow than *P. maniculatus* is a critical first step toward identifying the targets of natural selection. To address this challenge, we first developed a novel video-recorded assay to examine, specifically, the previously unexplored temporal components of burrowing behavior. We then characterize the genetic architecture of complex burrowing behavior by performing QTL analysis on each of these component traits.
5.3 Materials and Methods

Behavioral Assay. We recorded the overnight behavior of mice in an acrylic chamber (5cm x 0.9m x 0.6m) fitted with a transparent Plexiglas face through which the animal could be observed (Fig. 1b). The enclosure was outfitted with an infrared illuminator frame that enables visualization of the mouse’s nocturnal behavior using a Sony HD Camcorder equipped with an IR lens filter. Using a pre-cut mould, we sculpted hydrated sand into two symmetrical 45° hills and excavated a small artificial pre-dug tunnel into the side of one hill to elicit burrowing. Next, the mouse was introduced to the enclosure at the start of the dark cycle (16L:8D) and removed the following morning. We provided food and water ad libitum. Each mouse was tested twice (between 50 and 60 days old) to control for intra-individual variation in behavior. Photographs of the resultant burrows were taken at the end of the trial; details of photo and video analyses are provided below. Following each trial, we removed all food, feces, and disturbed substrate and rinsed the enclosure to minimize residual odours.

Video and Photo Analyses. We analysed 8 hours of overnight video per trial using custom Python scripts (code available at https://github.com/david-zwicker/cv-mouse-burrows). From this analysis, we extracted several new behaviors including various aspects of mouse movement and real-time burrow dynamics. To accurately measure the size and shape of each burrow, we complemented real-time burrow tracking with image analysis of still-photos taken at the end of each trial. We hand-traced the final burrow outline of up to two burrows per trial using FIJI image processing software (Schindelin et al. 2012). From these polygons, we extracted several two-dimensional components of burrow size and shape (see Table 1 for details).

QTL Mapping. Cross design and genotyping strategy follow Bendesky et al. 2017. In brief, an F2 mapping population of 568 males was generated by mating 4 P. maniculatus females to 4 P. polionotus males and subsequently crossing 16 pairs of first-generation (F1) hybrid full siblings. Flint
suggested that 600 individuals represents the upper bound of sample size required to detect behavioral QTLs in mice. All mice were genotyped using the ddRAD-seq/MSG approach described in Bendesky et al. 2017. Ultimately, 56,068 variants were used for QTL mapping. We performed genome scans by Haley-Knot regression on normalized traits in R/qtl (Broman et al. 2003). To assess genome-wide significance, we ran permutation tests with 1000 permutations for the autosomes and 16166 permutations for the X chromosome (Broman et al. 2006).

5.4 Results

Species-specific burrowing behavior is repeatable across multiple trials

First, to determine if burrowing behavior is repeatable, we tested 6 individuals of each species in 6 consecutive trials (Fig. 1c). Neither species showed a significant effect of trial number on burrow size (P. man, P = 0.466; P. pol, P = 0.572), indicating that burrowing ability neither improves nor declines over time.

Video analysis reveals previously unknown behavioral phenotypes

Next, to parse differences in burrowing behavior between P. maniculatus and P. polionotus, we compared several component traits derived from our automated video analysis (Fig. 1d, e). We failed to detect any difference in total time spent moving between the two species (P = 0.26). However, P. polionotus spent more time moving underground (P = 6.87e-9) and more time digging than P. maniculatus (P = 6.57e-4) (Fig. 1d). P. polionotus also began burrow construction earlier (P = 2.47e-3) and reached a peak digging rate earlier in the night than P. maniculatus (P = 9.78e-3) (Fig. 1e). Together, these results suggest that P. polionotus is not generally more active than P. maniculatus but, rather, that P. polionotus spends more of its active time underground. Differences in the timing of burrow
construction may further indicate that *P. polionotus* is more motivated to dig—starting sooner and completing most of its burrow construction earlier in the night than *P. maniculatus* (Fig. 1e).

![Image of burrow construction](image)

**Figure 5.1: Novel burrowing assay reveals key differences in burrowing behavior between two *Peromyscus* species.**

- **a**, Schematic of divergent burrow architectures between *P. maniculatus* (red) and *P. polionotus* (blue).
- **b**, Species-specific burrowing behavior is recapitulated during a video-recorded overnight assay in a transparent chamber.
- **c**, Burrowing behavior is repeatable across trials in *P. maniculatus* (left) and *P. polionotus* (right) (*n* = 6 *P*. man; *n* = 6 *P*. pol).
- **d–e**, Species distributions of selected behavior duration metrics (d) and behavior timing metrics (e) (*n* = 24 *P*. man; *n* = 56 *P*. pol).
- **f**, Principal Component Analysis (PCA) of burrow architecture attributes (left) and mouse behaviors (right).
- **g**, Relationship between time spent moving underground and burrow architecture PC1. Significance levels: ns = non-significant, **P < 0.01, ***P < 0.001.

To identify component traits that explain the most variance in burrowing behavior, we next per-
formed Principal Component Analysis (PCA). We ran this analysis separately for burrow architecture and mouse behavior traits (see Table 1). For burrow architecture, the first principal component (PC1) explained 52.7% of the total variance and its top-loading trait was longest burrow length (Fig. 1f, left). Strikingly, *P. maniculatus* and *P. polionotus* had largely non-overlapping PC1 values. By contrast, PC1 for mouse behavior did not clearly distinguish between the species. PC1 explained 40.6% of the total variance in mouse behavior and its top-loading trait was mouse mean speed (Fig. 1f, right). Thus, while the majority of variance in burrow architecture is largely partitioned by species, the majority of variance in mouse behavior is explained by a trait that does not clearly distinguish *P. maniculatus* from *P. polionotus*.

Additional methods are therefore required to identify the species differences in behavior that are associated with species differences in burrow architecture. Accordingly, we ran a Linear Discriminant Analysis (LDA) to determine which trait best explains variation in mouse behavior between *P. maniculatus* and *P. polionotus*. We found that the two species differed most strongly according to their time spent moving underground—a trait that correlates positively with PC1 for burrow architecture (Fig. 1g, R² = 0.726, P < 2e-16). These results suggest that, while the behavior that varies most among individuals is mean speed, the best behavioral predictor of species differences in burrow architecture is time spent moving underground.

**Trait correlations reveal additional species differences**

To determine if correlations between traits—in addition to the component traits themselves—differ between *P. maniculatus* and *P. polionotus*, we constructed trait correlation matrices for each species (Fig. 2a, b). Unsurprisingly, we found that some pairs of traits were nearly perfectly correlated, while others were highly anti-correlated. However, we found interesting species differences in trait correlations that shed light on the behavioral mechanism by which *P. polionotus* constructs longer burrows than *P. maniculatus*. 

95
Figure 5.2: Trait correlations reveal additional species differences in burrowing behavior. a. Pearson correlation coefficients for pairs of traits in P. maniculatus. Component trait order for P. polionotus is applied to P. maniculatus correlation matrix. b. Pearson correlation coefficients for pairs of traits in P. polionotus. Component traits are ordered according to hierarchical clustering based on Euclidean distance (see Fig. D1 for trait labels). c-e. Selected regression plots for pairs of traits in P. maniculatus (red, n = 24) and P. polionotus (blue, n = 56). c. Relationship between burrow initiation time and final burrow length. d. Relationship between peak digging time and final burrow length. e. Relationship between time spent digging and final burrow length.

In P. maniculatus, we found no significant associations between final burrow length and when mice began burrow construction (Fig. 2c, P = 0.146) or when they achieved their peak digging rate (Fig. 2d, P = 0.575). By contrast, within P. polionotus, mice that initiated burrow construction earlier (Fig. 2c, P = 1.36e-5) and reached their peak digging rate earlier in the trial (Fig. 2d, P = 3.46e-3) produced longer final burrows. In both species, time spent digging was positively correlated with
burrow length ($P.\ man, P = 2.61e^{-4}; P.\ pol, P = 3.81e^{-13}$) (Fig. 2e). However, the slope of the regression line differed between species ($P.\ man, m = 3.048; P.\ pol, m = 15.68$). Together, these findings indicate that motivated $P.\ polionotus$ individuals (i.e., those that start early and peak early) produce longer burrows. Furthermore, $P.\ polionotus$ digs approximately five times more efficiently than $P.\ maniculatus$. Thus, increased burrow length in $P.\ polionotus$ cannot be explained by an overall increase in general activity level. Rather, $P.\ polionotus$ achieve longer burrows by spending more time digging and through increased digging efficiency.

**QTL mapping identifies chromosomal regions involved in burrowing behavior**

With these behavioral mechanisms in mind, we first performed QTL analysis for a key component trait—time spent digging. To do so, we generated a mapping population of 568 F2 males. Each mouse was tested twice in the overnight assay and we performed QTL mapping with the mean trait value of trials 1 and 2. We found a QTLs for time spent digging on chromosomes 1, 21 and X. On average, $P.\ polionotus$ spends 68 minutes digging whereas $P.\ maniculatus$ spends only 40 minutes
We detected significant QTLs on chromosomes 1, 2, 3, 21, and X. Each row corresponds to a different component trait. We partitioned mouse behavior traits into mouse movement (pale violet) and burrow construction (pale green) behaviors. Burrow architecture traits are indicated in pale brown. Rectangles represent the boundaries of the Bayes credible interval for the QTL and are colored according to the LOD score for each genetic marker in the interval.

F2 mice that inherited at least one *P. polionotus* allele at the genetic marker corresponding to the QTL peak on chromosome 1 spent, on average, an additional 11.5 minutes digging.
We next performed QTL analysis for all component traits (Fig. 4). We observed a number of interesting trends in these data. First, burrow architecture traits mapped exclusively chromosomes 1 and X. Second, mouse movement traits—such as speed, time spent moving, and distance covered—mapped to chromosomes 2, 3, and 21. Third, mouse behavior traits associated with burrow construction—such as time spent underground and time spent digging—mapped to chromosomes 1, 21, and X. Surprisingly, all QTLs detected on chromosomes 21 and X were transgressive (i.e., the allelic effects were opposite those predicted from parental phenotypes). For instance, F2s that inherited a P. poloionotus allele at the genetic marker corresponding to the X chromosome QTL peak for time spent digging actually spent less time digging than F2s that inherited a P. maniculatus allele at this locus.

5.5 Discussion

A major challenge for genetic mapping of complex behavior is developing a precise phenotyping assay that consistently elicits full expression of the behavior of interest. From such an assay, reliable, repeatable, and biologically relevant measures can be made, ideally in an unbiased and automated fashion. Previously, we circumvented this challenge by simply measuring the final product of behavior—the burrows themselves. Using this approach, Weber et al. (2013) identified three independent genetic regions associated with variation in entrance tunnel length and one locus associated with the presence of an escape tunnel. While easily quantifiable, these simple measures failed to capture the many ways in which a complex burrow structure might be achieved. For instance, an individual may devote more time to burrow construction, or it may dig at a faster rate. Both behaviors could result in increased burrow length, yet these alternate strategies are not clearly discerned by simply measuring the final outcome of several behaviors integrated over the course of the night (i.e., burrow structure).
Previous studies indicate that the motor patterns employed by \textit{P. maniculatus} and \textit{P. polionotus} largely overlap. Comparative work on the digging habits of muroid rodents concluded that both \textit{P. maniculatus} and \textit{P. polionotus} simultaneously kick with both front and rear paws, rather than alternating kicks like most other muroid rodents (Layne and Ehrhart 1970, Webster et al. 1981). Thus, while basic motor patterns appear conserved within \textit{Peromyscus}, our study reveals several striking interspecific differences in the temporal components of burrow construction that were previously unobservable. Despite equivalent levels of overall activity, we found that \textit{P. polionotus} was more active underground, spent twice as much time digging, and was five times more efficient at digging as \textit{P. maniculatus}. These findings raise the possibility that complex burrowing behavior in \textit{P. polionotus} arose by genetic changes that affect motivational states in the brain, leading \textit{P. polionotus} to dig early and often. To isolate these genetic changes, we employed quantitative trait locus (QTL) mapping in F2 hybrids.

QTL mapping is a powerful tool for identifying loci that underlie morphological, life-history, and behavioral traits (Gutteling et al. 2007, Mackay et al. 2009, Yalcin et al. 2004). However, there is a large disconnect between locating QTL and nominating candidate genes. Thus far, the identification of a great many mouse QTL (2000) has not translated into an equally large number of causal genes (20) (Flint et al. 2005). Here, we used a QTL mapping approach to locate the genetic regions and estimate the effect sizes of loci underlying variation in several components of burrowing behavior. We identified QTL for 28 component traits on 5 chromosomes, with effect sizes ranging from 3.5 to 11%. Although QTL mapping is an unbiased, genome-wide, and direct approach that narrows and localizes the pool of potential candidate loci, these QTL peaks still span large regions of the genome (10 to >100 Mb) and harbour many genes. Gene density in the mouse genome can range from tens to hundreds of essential genes per 20 Mb interval (Hentges et al. 2007). This is far too great a number to functionally test each gene in the QTL in an unbiased manner.

Alternative approaches are therefore required to narrow the QTL peaks identified here and nom-
inate candidate genes. Future efforts could compare the gene expression profiles of pure species brains, using massively parallel RNA sequencing (RNAseq). Furthermore, RNAseq could be used to quantify allele specific expression (ASE) in the brains of first-generation (F1) hybrids to identify changes in cis-regulatory elements that may contribute to species-specific differences in burrowing behavior. The most promising candidate genes will be those that emerge at the intersection of genetic mapping and gene expression profiling analyses. A gene or genetic element that is situated under a QTL peak, that is differentially expressed in the brain between *P. maniculatus* and *P. polionotus*, and whose expression is controlled by a cis-acting polymorphism, is a strong candidate for functional validation. Together, this work lays the foundation for the functional verification of candidate genes and offers a novel mechanistic understanding of the precise behavioral changes that drive complex trait evolution.

5.6 References


A comprehensive understanding of behavioral diversity requires scientific inquiry from many angles. Here, I have applied Tinbergen’s four questions to better understand variation in burrowing behavior among *Peromyscus* mice. By applying complementary methods from animal behavior, quantitative genetics, and behavioral ecology, I put forward a more complete understanding of how and why behavioral variation evolves. In sum, these different levels of analysis advance the comparative ethology of burrowing behavior in *Peromyscus*. 
Supplementary Material for Chapter 2
A.1 Materials and Methods

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 (≥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 5 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 5 for measurements of body mass for this species across development.
Cross-fostered mice. Age-matched (≤ 48 hrs age difference) *P. maniculatus* (n=18; 8 females, 10 males) and *P. polionotus* (n=16; 5 females, 11 males) pups were traded between experienced (≥1 previous litter) heterospecific breeding pairs 24-48 hours 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* F1 Hybrids. We produced F1 hybrids by crossing *P. maniculatus* dams to *P. polionotus* sires. Due to genomic imprinting in these species, our cross design for production of F1 hybrids was limited to one direction [49,50]. 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 F1 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 F1 hybrids to *P. maniculatus* mates (Figure 3a). Both male and female F1 hybrids were backcrossed to *P. maniculatus* (reciprocal pairings); 22 animals were produced from an F1 dam and 38 from an F1 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 F1 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 hours to 14-17 hours (i.e. one 8-hour dark cycle followed by 6-9 hours 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 Corp.), 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 F1 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 [51]. We tested naïve, 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 hours of home cage habituation to the wheel (Ware Manufacturing Inc., Phoenix, AZ), we recorded 90 minutes of wheel running activity with a CC-COM10W wireless bike computer (Cateye Co. Ltd., Osaka, Japan). Peromyscus show strongly nocturnal patterns of wheel running [51], thus we performed all tests during the first 4 hours 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 [5]) 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 A2). 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 F1 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 A2), 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 Analyses**

**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 ≥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 Corp., 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.

**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 [53] 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 [54], and p-values for all fixed effects were calculated using mixed afex [55].

**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 [53] 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 [54], and p-values for fixed effects were calculated using mixed afex [55]. To further evaluate the significance of these genotype-phenotype associations, we implemented Benjamini-Hochberg corrections [52] 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 [56]. 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 [8].
Figure A.1: Comparison of morphological and motor development in *P. polionotus* (blue) and *P. maniculatus* (yellow) juveniles. (A) Morphological growth as a percentage of adult body mass, body length and hindfoot length (data adapted from [S1]), as well as absolute body mass, absolute hindfoot length, and distance running on a wheel. (B) Ontogenetic trajectory of *P. polionotus* and *P. maniculatus* at key developmental events (data adapted from [S1, S2] as indicated).
Figure A.2: Effect of genotype at four markers [S3] on juvenile and adult burrowing behavior in BC hybrids. Three loci were previously associated with adult tunnel length: (A) Marker 1 (Atxn2); (B) Marker 2 (Chrm5); and (C) Marker 3 (Ntrk3). One locus was associated previously with adult escape tunnel construction: (D) Marker 4 (Ndrg4). Genotypes for BC hybrids are either MM (homozygous for the *P. maniculatus* allele) or PM (heterozygous). For each genotype, trait means are plotted with error bars at ± 1 SE of the mean. Significance levels, determined by linear mixed-effect models with repeated measures and Benjamini-Hochberg corrections with 10% FDR are: $p \leq 0.05 = *$, $p \geq 0.05 =$ not significant (ns).


Table A.1: Related to Chapter 2 Figure 1. Proportion of mice digging complete burrows (i.e. entrance tunnel and nest chamber) at three stages: pre-weaning (P19-25), weaned juveniles (P27-31) and adults (≥P60). Significance levels for p-values (Fisher’s exact tests): p ≤ 0.05 = *, p ≤ 0.01 = **, p ≤ 0.001 = ***.

<table>
<thead>
<tr>
<th></th>
<th>P. pol 19-25 days (n = 32)</th>
<th>P. pol 27-31 days (n = 21)</th>
<th>P. man 19-25 days (n = 19)</th>
<th>P. man 27-31 days (n = 20)</th>
<th>P. man Adult (n = 17)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P. pol Adult</td>
<td>p = 0.071</td>
<td>p = 1.000</td>
<td>p &lt; 0.0001 ***</td>
<td>p &lt; 0.001 ***</td>
<td>p = 0.023 *</td>
</tr>
<tr>
<td>P. pol 27-31 days</td>
<td>p = 0.165</td>
<td>p &lt; 0.0001 ***</td>
<td>p &lt; 0.001 ***</td>
<td>p = 0.005 **</td>
<td></td>
</tr>
<tr>
<td>P. pol 19-25 days</td>
<td>p = 0.047</td>
<td>p &lt; 0.001 ***</td>
<td>p = 0.101</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P. man Adult</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P. man 27-31 days</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P. man 19-25 days</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P. man 27-31 days</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table A.2: Related to Chapter 2 Figure 4. Primer sequences and restriction enzymes used for RFLP genotyping assays at four loci. (frag = fragment length).

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Primer direction</th>
<th>Primer sequence (5’ to 3’)</th>
<th>SNP location</th>
<th>Restriction enzyme (cut site)</th>
<th>P. man frag (bp)</th>
<th>P. pol frag (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ntrk3</td>
<td>F</td>
<td>TGATCCCAAGAAC CAGGATCACA</td>
<td>4,884,888</td>
<td>HinfI (GAATC)</td>
<td>178</td>
<td>73 + 105</td>
</tr>
<tr>
<td>Ntrk3</td>
<td>R</td>
<td>CCAAAGGGAAAAA CAAACCTAAAAGG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chrm5</td>
<td>F</td>
<td>GTTCTTCAAAAC AAGCAAACCAGAC</td>
<td>378,196</td>
<td>HinfI (GAATC)</td>
<td>381</td>
<td>123 + 257</td>
</tr>
<tr>
<td>Chrm5</td>
<td>R</td>
<td>CAAGCGAAAGAG AATGGTCTCTAATC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atxn2</td>
<td>F</td>
<td>ATTTTTCCAATA TAGTCTGGCAGT</td>
<td>6,163,590</td>
<td>HinfI (GAATC)</td>
<td>245</td>
<td>97 + 147</td>
</tr>
<tr>
<td>Atxn2</td>
<td>R</td>
<td>ATCTTCAGGGGT TATGTATCCCAAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ndrg4</td>
<td>F</td>
<td>GCCCAGCTTCCC ACAATGAT</td>
<td>2,219,511</td>
<td>MseI (TTAA)</td>
<td>406</td>
<td>216 + 190</td>
</tr>
<tr>
<td>Ndrg4</td>
<td>R</td>
<td>CTCCCAAGCGACG ATCAAGGT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Supplementary Material for Chapter 3
B.1 Materials and Methods

**Animal husbandry**

We performed experiments with *P. leucopus* (LL stock), *P. maniculatus bairdii* (BW stock), and *P. polionotus subgriseus* (PO stock) maintained as outbred colonies at Harvard University. Stocks were originally obtained from the Peromyscus Genetic Stock Center at the University of South Carolina. We maintained animals in ventilated cages measuring 7.75 x 12 x 6.5” (Allentown Inc., Allentown, NJ), which were furnished with 1/4” Bed-o’Cobs bedding (The Andersons, Maumee, OH), a 2” square cotton nestlet (Ancare, Bellmore, NY), and a red polycarbonate hut (Bio-Serv, Flemington, NJ). We provided food and water ad libitum. Upon weaning at 23 days, mice were co-housed in same-sex groups of at most five conspecifics and fed irradiated Prolab Isopro RMH 3000 (LabDiet, St. Louis, MO). We fed all paired adults irradiated breeder chow: PicoLab Mouse Diet 20 (LabDiet, St. Louis, MO). We maintained mice on a 16h light:8h dark cycle at 22°C. The Institutional Animal Care and Use Committee at Harvard University approved all protocols.

**Behavioral assays**

Large opaque enclosures. We measured burrow architecture as described previously. We measured burrow architecture as described previously.12,46 Briefly, we introduced individuals or pairs of mice into large PVC boxes (1.2 x 1.5 x 1.1m) filled with 700kg hydrated, hard-packed premium play sand (Quikrete, Atlanta, GA). We provided food and water ad libitum along with a cotton nestlet. We removed mice from the enclosures after one (pregnancy trials) or two (standard trials) overnight active periods. Then, we made casts of the resultant burrows using polyurethane filling foam (Hilti, Schaan, Liechtenstein). Next, a researcher blind to trial identity hand-measured burrow casts. After each trial, we removed all food, feces, nesting material, and disturbed substrate from the enclosure, and then rinsed the enclosure walls with water and
turned over the sand to minimize residual odours.

**Narrow transparent enclosures.** We also assayed the burrowing behaviour of mice directly using an acrylic chamber (5cm x 0.9m x 0.6m) with a transparent Plexiglass face (see Fig. 3a). Using a pre-cut mould, we sculpted hydrated sand into two symmetrical 45° hills and excavated an 8cm starter tunnel from one randomly selected hill to encourage burrowing in a consistent location. We outfitted the apparatus with an infrared illuminator frame that enabled video recording in the dark. We then introduced individuals or pairs of mice into the enclosures at the start of the dark cycle and removed animals the next morning, recording 8 hours of continuous video during the dark cycle. We provided food and water ad libitum. Photographs of the resultant burrows were taken at the end of the trial; details of photo and video analyses are provided below. Following each trial, we removed all food, feces, and disturbed substrate and rinsed the enclosure to minimize residual odours.

**Experimental design**

**Individual and pair burrowing differences across species.** We first measured the burrows of the three *Peromyscus* species in individual and pair trials in the large opaque enclosures. To control for past experience, we varied the order in which mice were tested in individual, same-sex, or opposite-sex trials. On average, each individual and each unique pair were tested twice. Mice were released into the enclosures at the start of the dark cycle and retrieved after two full overnight periods (36 hours).

**The effect of pregnancy on female burrowing.** To test whether pregnancy alters burrowing performance, we compared the burrows dug by females before and after being co-housed with a male. First, we tested the burrowing output of virgin females (n = 42 *P. polionotus*, n = 36 *P. maniculatus*) in the large enclosures. Each female was assayed twice with 2 days rest between trials. We then transferred each female to a new cage with a conspecific male. Approximately half of all
paired females subsequently became pregnant (21/42 *P. polionotus*, 19/36 *P. maniculatus*). Co-housed females were then tested again twice—as individuals—with 2 days rest between trials. Females were returned to their home cage after each trial and monitored daily for parturition for up to 23 days. To minimize stress to pregnant females, all mice in this experiment were tested for one overnight period (18 hours) in the large opaque enclosures.

**Social modulation of burrowing behaviour in *P. polionotus***. To test whether *P. polionotus* mice modulate their behaviour in response to a digging partner, we used a repeated measures design. We video-recorded all trials in the narrow transparent enclosures (see Fig. 3a). To distinguish individuals, we shaved a patch of hair from both flanks of one randomly-selected member of the pair. Shaving was completed at least a week before the first pairing, and markings did not grow in over the course of the experiment. To quantify baseline individual burrow output, we first assayed each individual twice (*n = 16* virgin females, *n = 16* virgin males). Next, we transferred each individual to a new cage with an unrelated and unfamiliar partner of the same or opposite sex. Pairs were given 2 nights to acclimate before being tested together in the small transparent enclosures. We recorded 8 hours of video for each overnight trial. Pairs were then returned to their home cage for 3 nights rest before their second trial. Mice were then re-partnered and the process was repeated. To control for previous experience, we randomly assigned the order in which mice burrowed with a same-sex or opposite-sex partner. Females were monitored daily for parturition for up to 23 days after the end of the experiment. No females were pregnant at any point during the experiment.

**Behavioural Analyses**

**Photo analysis.** We measured the length of burrows produced by individuals or pairs in the narrow enclosure using Fiji image processing software. Each photo was measured separately by two researchers blind to trial identity. Burrow length measurements were highly correlated between researchers (Pearson correlation, *r = 0.97, P = 2.2e-16*). We therefore used the mean of two measure-
ments for all statistical analyses. To calculate digging efficiency, we also took video snapshots at the beginning and end of each 10-minute observation timepoint. From these snapshots, we measured burrow length as described above.

**Video Analysis.** In 75% of trials, pairs had fully completed at least one burrow 3 hours after lights out. Thus, to target the active burrow-construction phase, we selected two 10-minute timepoints, spaced 1 hour apart, in this 3-hour window for observation: 0:30–0:40 and 1:30–1:40 hours. In 5 of the 62 videos, we did not observe digging during the pre-selected timepoints. For these videos, we shifted our two timepoints, still spaced one hour apart, to later in the night. We accounted for this shift in all statistical models.

We quantified the behaviour of each mouse in a pair using The Observer XT Version 12.0 (Noldus, Leesburg, VA). Each 10-minute timepoint was randomly assigned to a researcher blind to trial identity for behavioural scoring according to the following scheme. For each individual, we scored all burrow entries or exits, as well as extending (i.e., forelimb digging at the growing end of the burrow), widening (i.e., forelimb digging at any other location in the burrow), and hind-kicking (i.e., vigorous, coordinated hindlimb movements that expel loosened sand from the burrow). For each pair, we also scored affiliative and agonistic behaviours, which could occur either inside or outside the burrow. We defined affiliative behaviours as beneficial social interactions (i.e., allogrooming, huddling). We defined agonistic behaviours as aggressive interactions (i.e., boxing, parrying, biting) or submissive behaviours (i.e., freezing, fleeing). Altogether, each mouse \( n = 32 \) received 80 minutes of direct observation across the entire study—10 minutes per timepoint, 2 timepoints per trial, 2 trials per trial-type, 2 trial-types (same-sex and opposite-sex). We excluded observations from one same-sex pair in which the markings of the two females were indistinguishable and ID could not be confidently assigned.
**Statistics**

All statistical tests were performed in R Version 3.2.3 (R Core Team, 2015). **Individual and pair burrowing differences across species.** To normalize the data, we first square-root transformed burrow lengths. To test if pairs of mice dig longer burrows than individuals, we used a linear mixed-effects model (LMM) with social context (i.e., individual or pair trial) and sex as fixed effects and mouse ID as a random effect. We excluded trials in which individuals or pairs did not burrow. For each species, we divided the length of the average burrow dug by two mice by the length of the average burrow dug by one mouse to determine the pair: individual burrow length ratio (n = 9 pairs, 15 singles (P. leu); 22 pairs, 21 singles (P. man); 52 pairs, 52 singles (P. pol) (Fig. 1b). We approximated the variance of each ratio by Taylor series expansion. To further evaluate sex and pair-type differences within species, we used LMMs with either sex or pair-type as a fixed effect and mouse ID as a random effect (Fig. 1c). Differences in observed: expected burrow length ratio between same-sex and opposite-sex pairs were evaluated by LMM, with mouse ID and partner ID as random effects (Fig. 1d).

**The effect of pregnancy on female burrowing.** Data were log-transformed to improve normality. To test if individual females dig longer burrows after being co-housed with a male, we used LMMs with stage (i.e., before or after male-cohabitation), status (i.e., non-pregnant or pregnant), and their interaction as fixed effects and mouse ID, age, and length of cohabitation period as random effects. The non-significant interaction term was subsequently removed from the *P. polionotus* model. For *P. maniculatus*, we used a least-squares means post-hoc test to further evaluate the stage by status interaction. For both species, we used binary logistic regression to test if females with longer virgin burrows were more likely to become pregnant.

**Social modulation of burrowing behaviour in *P. polionotus.*** When necessary, data were transformed to improve normality. To test the effect of social context on burrow length, we used LMMs with trial-type (i.e., individual, same-sex, or opposite-sex) as a fixed effect. We included
mouse ID and trial number as random effects to control for repeated measures and whether mice first burrowed with a same-sex or opposite-sex partner.

**Sex differences.** Because we found no sex or trial-type differences in the proportion of burrow extending vs. burrow widening behaviour (LMM, sex: $P = 0.490$, trial-type: $P = 0.384$), we collapsed these two categories into one general “digging” designation. To assess overall sex differences in digging behaviour, we used LMMs with mouse sex and trial-type (i.e., same-sex or opposite-sex) as fixed effects and mouse ID, observer ID, trial number, and timepoint as random effects. To test explicitly for social modulation of digging behaviour, we used LMMs with mouse sex, trial-type, and their interaction as fixed effects and mouse ID, observer ID, trial number, and timepoint as random effects. For the underground bout model, we used a least-squares means post-hoc test to further evaluate the sex by trial-type interaction.

**Pair-type differences.** To calculate division of labour, we divided the total digging duration for each mouse by the total digging duration for the pair. Individuals in same-sex pairs were distinguished by their markings (i.e., shaved or non-shaved) and individuals in opposite-sex pairs were distinguished by sex. We used LMMs to ask if individuals comprising a pair differed in their contributions, including mouse ID, observer ID, trial number, and timepoint as random effects (Fig. S1a). We used LMMs and Tukey’s post-hoc comparisons to evaluate differences in total digging duration and total burrow length among pair-types (Fig. S1b, c). For total digging duration, we included pair ID, observer ID, trial number, and timepoint as random effects in the LMM. For total burrow length, we included pair ID and trial number as random effects.

**Differences in social behaviour.** To quantify differences in social interaction among pair-types, we summed the number of affiliative encounters per pair across all observations. We then used an LMM with pair-type as a fixed effect and pair ID, observer ID, trial number, and timepoint as random effects. Since Tukey's test revealed no significant difference between FF and MM ($P = 0.726$), we re-ran the above LMM with trial-type (i.e., same-sex or opposite-sex) as a fixed effect.
rather than pair-type (i.e., FF, MM, or FM). Agonistic interactions were much rarer. We therefore
counted the number of trials with and without any observed agonistic interactions and used pair-
wise Fisher's exact tests to evaluate differences among pair-types.

**Modelling burrowing efficiency.** We modelled the relationship between independent
digging, simultaneous digging, and change in burrow length using multiple linear regression. We
then used binary logistic regression to test if opposite-sex pairs were more likely than same-sex pairs
to engage in simultaneous digging, controlling for the time spent digging by both the focal mouse
and its partner.
Figure B.1: Different pair configurations differ in total time spent digging, burrow length, and division of labour.

a, Division of labour in same-sex and opposite-sex pairs. On average, each mouse in a same-sex pair contributed equally to the total time spent digging (LMM, P = 0.096, n = 30), whereas the division of labour in opposite-sex pairs was significantly skewed (LMM, P = 1.68e-7, n = 31). Data points represent the mean of 4 observations per individual, per trial-type. Within same-sex trials, grey points on the left represent shaved mice (S) and points on the right represent non-shaved mice (N). b, Total time spent digging by both mice in a pair, per 10-minute observation period. Data points represent the mean of 4 observations per pair. Opposite-sex pairs and male pairs spent more total time digging than female pairs (Tukey’s test, FM vs. FF: P = 0.039; MM vs. FF: P = 0.004, FM vs. MM: P = 0.406). c, Total burrow length produced by different pair configurations. Data points represent the mean of 2 trials per pair. Opposite-sex pairs and male pairs built longer burrows than female pairs (Tukey’s test, FM vs. FF: P = 0.003; MM vs. FF: P = 0.006; FM vs. MM: P = 0.491). d-f, Summary statistics for behaviours quantified in 10-minute observation periods. Data points represent the mean of 8 observations per individual. Behaviour labels are as follows: “in burrow” (i.e., mouse is underground), “extending” (i.e., mouse is digging at the leading edge of the burrow), “widening” (i.e., mouse is digging at the centre of the burrow), “obscured” (i.e., mouse is underground but not visible), “hindkick” (i.e., powerful expulsive kick that ejects loose sand from the burrow). d, Total duration of behaviour, per 10-minute observation period. e, Number of bouts, per observation. f, Mean bout duration, per observation. Significance levels: *P < 0.05, **P < 0.01, ***P < 0.001. Error bars represent SEM.
Supplementary Material for Chapter 4
C.1 Materials and Methods for November 2017 Fieldwork

We collected and analysed data from two field seasons, April 30th – May 18th 2016 and November 20th – December 1st 2017. We employed the same experimental design and analytical approaches in both years. Results from November 2017 are reported in the main text and are very similar to the May 2016 results, which are reported in the supplement (Appendix C).

Phase I: Capture-recapture (November 20th – November 18th 2017)
We established grids of Sherman folding traps (H.B. Sherman Traps, Inc., Tallahassee, FL) in the frontal dunes of Santa Rosa Island, FL (30°23’42” N, 86°43’27” W). We set a 70 x 9 trap grid, with traps spaced at 10m intervals along both east-west and north-south axes. The final grid measured 690 x 80m ($55,200 \text{ m}^2$) and contained 630 trap locations. We trapped for 16 consecutive nights, but not all traps were set on all nights; see Table C1 for usage information. Following the grid trapping period, we set 3-4 traps near the entrances of putative burrow sites for an additional two nights. Using this method, we trapped 7 new animals (1 adult and 6 juveniles, 2 of which were too small to tag). We did not include these additional captures in our spatially explicit capture-recapture (SECR) models. Altogether, we trapped 43 mice, 39 of which we successfully tagged and released.

We set traps before sunset each night and checked at sunrise the following morning. Trapped beach mice were sexed, weighed, measured, and injected with an 8mm, 125 kHz RFID tag (Freevision Technologies Co., Shanghai, China). We also sampled a small piece of ear tissue for DNA extraction. We then returned each mouse to its original trap location and noted if the animal entered a burrow upon release.

Phase II: RFID monitoring (November 20th – December 1th 2017)
We identified putative burrow sites by observing which burrows mice entered following release.
Table C.1: November 2017 capture-recapture summary.

<table>
<thead>
<tr>
<th>Night</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traps set</td>
<td>80</td>
<td>328</td>
<td>414</td>
<td>414</td>
<td>414</td>
<td>495</td>
<td>495</td>
<td>495</td>
</tr>
<tr>
<td>Captures</td>
<td>4</td>
<td>4</td>
<td>7</td>
<td>8</td>
<td>3</td>
<td>7</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>New captures</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Losses</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cumulative</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>9</td>
<td>11</td>
<td>11</td>
<td>12</td>
<td>16</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Night</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traps set</td>
<td>495</td>
<td>495</td>
<td>495</td>
<td>495</td>
<td>495</td>
<td>243</td>
<td>135</td>
<td>6483</td>
<td></td>
</tr>
<tr>
<td>Captures</td>
<td>6</td>
<td>10</td>
<td>10</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>6</td>
<td>5</td>
<td>128</td>
</tr>
<tr>
<td>New captures</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>36</td>
</tr>
<tr>
<td>Losses</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Cumulative</td>
<td>21</td>
<td>25</td>
<td>26</td>
<td>29</td>
<td>33</td>
<td>34</td>
<td>36</td>
<td>36</td>
<td>36</td>
</tr>
</tbody>
</table>

Table C.2: November 2017 RFID summary.

<table>
<thead>
<tr>
<th>Night</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active RFID readers</td>
<td>26</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>28</td>
<td>28</td>
<td>23</td>
<td>22</td>
<td>22</td>
<td>40</td>
</tr>
<tr>
<td>No. mice detected</td>
<td>20</td>
<td>21</td>
<td>18</td>
<td>20</td>
<td>23</td>
<td>21</td>
<td>20</td>
<td>18</td>
<td>21</td>
<td>20</td>
<td>17</td>
<td>32</td>
</tr>
<tr>
<td>Total activity (no. pings)</td>
<td>1152</td>
<td>1003</td>
<td>797</td>
<td>783</td>
<td>1137</td>
<td>876</td>
<td>621</td>
<td>382</td>
<td>676</td>
<td>545</td>
<td>338</td>
<td>8310</td>
</tr>
</tbody>
</table>

Ghost crabs (*Ocypode quadrata*) are abundant in the foreshore and produce burrows similar in morphology to those of beach mice. Active beach mouse burrows were therefore distinguished from crab burrows—often used as temporary refugia—by trial and error. We then installed custom-built RFID readers at putative burrow entrances and kept readers in place for 2 consecutive nights (minimum). If no activity was recorded after 2 nights, the RFID reader was moved to a new putative burrow site. We recorded activity at a total of 40 unique burrows. Not all burrows were active on all nights; see Table C.2 for details. We observed RFID activity for 82% (32/39) of the mice that were tagged and released in Phase I.

Quantification and Statistical Analyses
Table C.3: November 2017 spatially explicit capture-recapture (SECR) model results.

<table>
<thead>
<tr>
<th>Value</th>
<th>estimate</th>
<th>s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prob. of capture at activity centre (go)</td>
<td>0.019</td>
<td>0.003</td>
</tr>
<tr>
<td>Scale parameter (s)</td>
<td>29.97</td>
<td>2.46</td>
</tr>
<tr>
<td>95% home-range radius (m)</td>
<td>67.26</td>
<td>-</td>
</tr>
<tr>
<td>Home-range size (m²)</td>
<td>14,212</td>
<td>-</td>
</tr>
<tr>
<td>Density (mice/km²)</td>
<td>254</td>
<td>48</td>
</tr>
</tbody>
</table>

SECR model. To estimate home-range size and population density, we ran a spatially explicit capture-recapture analysis using the secr package in R\textsuperscript{24}. In our model, we included all captures from nights 3-14 when > 400 traps were set. We used the “suggest.buffer” function to determine a suitable buffer width (80m) around the 690 x 80m trap grid. We then fit a half-normal spatial detection model by maximum likelihood using the “secr.fit” function, which estimates two parameters, go and \( \sigma \) (Table C3). The capture probability at the home-range centre is denoted by the intercept (go), and the scale parameter (\( \sigma \)) describes how this probability decays with increasing distance from the home-range centre. We estimated the 95% home-range radius from \( \sigma \) using the “circular.r” function.

Utilization distributions. We independently derived a second home-range size estimate using the spatial coordinates of RFID relocations rather than capture-recapture data with the function “kernelUD” from the adehabitatHR package in R\textsuperscript{25}. This function estimates the utilization distribution, a bivariate probability density function that gives the probability of relocating an animal at a given spatial coordinate. We included animals (\( n = 20 \)) that were detected on 7 or more nights, at 5 or more burrows, and that had 40 or more RFID pings over the course of the study (Fig. C1).

Burrow visitation patterns. We limited the following analyses to mice with 4 or more pings and burrows with 6 or more pings over the entirety of the 11-night study. In total, we analysed data from 26 mice and 36 burrows. To identify which burrows showed similar visitation patterns,
Figure C.1: Beach mouse home-range sizes estimated from utilization distributions. Utilization distributions and 95% home-range contours are calculated from spatial coordinates of RFID relocations. Both are drawn on the 690 x 80m trap grid. The sex, age, and 95% home-range size are indicated in the bottom-left corner of each panel (n = 20 mice). Data from November 2017.
we first determined the portion of each burrow’s total RFID activity comprised by different mice. 
We then scaled and centered these data and calculated a dissimilarity matrix based on Euclidean dis-
tance. We then used agglomerative hierarchical clustering to produce a cluster dendrogram of bur-
rows visited by similar groups of mice. We next defined 7 burrow clusters by drawing a horizontal 
line through the dendrogram at a minimum height such that all burrows within a cluster shared at 
least one mouse. Mice frequented burrows belonging to 1-3 clusters (mean = 1.65).

Spatiotemporal overlap. To determine the extent of spatiotemporal overlap between all 
pairs of mice in the RFID-tracked population, we developed an overlap index (OI). We defined 
OI_{i,j} as the portion of total RFID pings for mouse i that occur at the same burrow and within the 
same time-bin as mouse j. Because it is expressed as a portion of the focal mouse’s total activity, OI 
is not symmetrical (i.e., OI_{1,2} \neq OI_{2,1}). We used the “SlidingWindow” function from the evobiR 
package in R\textsuperscript{26} to calculate OIs for a range of time-bins, from 15 minutes to 14 hours, using a step 
size of \( \frac{1}{2} \) the time-bin size. For each time-bin, we then calculated the network load\textsuperscript{27}, defined as 
the number of observed links in the network divided by the total number of possible links (i.e., the 
fraction of all possible pair combinations with non-zero OIs). We selected an appropriate time-bin 
for determining spatiotemporal overlap by plotting network load as a function of time-bin size. We 
approximated 1000 points along this curve using the function “approx” in R and pinpointed the 
elbow of this curve (89 minutes) by calculating the second derivative. We therefore performed all 
downstream analyses on OIs calculated with a 1.5-hour time-bin.

Genetic relatedness. To identify polymorphic markers for \textit{P. \textit{p. leucocephalus}}, we re-analyzed 
the gDNA data of representative individuals available from a previous study by Domingues et al.\textsuperscript{28} 
(Short Read Archive accession number SRP010898). This dataset comprised data from 20 individu-
als sampled at 5000 1.5kb targeted non-coding regions randomly distributed across the genome (see 
aforementioned study for details about the capture array design). We aligned raw gDNA paired-end 
reads to an in-house assembly of the \textit{Peromyscus polionotus} genome (Genbank assembly accession:
GCA_003704135.1) using BWA (version 0.7.12)\textsuperscript{19}. The resulting alignment files were pre-processed according to the GATK Best practices and used to perform SNP and INDEL discovery and genotyping across all 20 samples separately using HaplotypeCaller as implemented in GATK version 3.\textsuperscript{7}\textsuperscript{31}. Default parameters were used, with the exception of the prior for heterozygosity which was set to 0.005. We then performed joint genotyping on the gVCF files produced by HaplotypeCaller and filtered the resulting call set using BCFtools version 1.7\textsuperscript{32} with standard hard filtering parameters (SNP: QD < 2.0, FS > 60.0, MQ < 40.0, SOR > 3.0, MQRankSum < -12.5, ReadPosRankSum < -8.0; INDEL: QD < 2.0, FS > 200.0, ReadPosRankSum < -20.0). In addition, we filtered out variants called at GQ<20 with a DP<7, excluded sites with missing genotype in more than 10 individuals, and selected only biallelic variant sites. The final set of polymorphic sites were identified using four criteria. 1) A site is considered variable if at least one sample of each category (hom_ref, hom_alt, het) is present. In addition, not all samples should be homozygous for either reference or alternate allele (1 < AC < 35). 2) Sites should not have an abnormally high number of heterozygous genotypes, as an excessive number of heterozygotes may reflect technical errors. We therefore excluded variants with ExcessHet > 7. 3) To identify the most variable sites, we calculated a ‘variability score’ defined as the product of the number of samples represented in each genotype category. For a given site, the higher the score, the more mixed the genotype profile is. 4) To avoid variants that were on the same segment, we excluded variants that were less than 200bp away from the nearest neighbor.

The resulting set of variants passing these filters comprised 294 sites. We extracted from the genome the nucleotide sequence corresponding to 100bp upstream and downstream of the 140 most variable sites (ranked based on our custom metric). We excluded markers that had more than one hit in the genome, as revealed by a local search using BLAST as implemented in Biopython\textsuperscript{13}. A final set of 110 markers was selected and used for the design of genotyping probes.

We then designed PCR primers for these target regions using BatchPrimer3 v1.0.
(https://probes.pw.usda.gov/batchprimer3/). We used the default settings for general primers and selected final amplicon sizes between 80 to 90bp. Multiplex amplification of the targets was done by Floodlight Genomics, LLC (Knoxville, TN) to produce sample-specific amplicons using an optimized Hi-Plex approach as part of a no-cost Educational and Research Outreach Program. Pooled barcoded amplicons were sequenced on a HiSeqX device (Admera Health, New Jersey US) and the sample-specific sequences were aligned to the target sequences used for primer design with CLC Genomics Workbench. Genotypes were assigned for loci with >10X coverage and an alternate allele frequency cut-off of 15% was used to assign heterozygous calls.

**Social network analysis.** We constructed network graphs using the igraph package in R. We calculated degree centrality for each node in the network using the “centr_degree” function.

**Home burrow determination.** We restricted our analysis to mice that were detected on all 11 nights of the RFID-tracking study. To determine where mice slept during the day, we extracted the first and last RFID ping of each night-time activity period. For each animal, we asked if the first ping at dusk was recorded at the same burrow as the last ping the previous dawn, indicating that the mouse had spent the day in that burrow. At the midpoint of the study, there was 10 hours and 26 minutes of daylight on Santa Rosa Island, with sunset and sunrise occurring at 4:46 PM and 6:20 AM, respectively. We found that mice spent between 11.3 and 18.8 hours in their burrows during the day (mean = 12.57 ± 0.15 s.e.).

**Burrow site attributes and temperature profiles.** At the end of the study we measured various attributes of the RFID-monitored burrows. To measure the length of the entrance tunnel, we inserted flexible plastic tubing into the burrow and recorded this distance. We also recorded the slope, elevation, and heading of the dune at the burrow entrance. Finally, to determine the extent of vegetative cover at the burrow site, we photographed the burrow entrance with a 50 x 50cm quadrat. Percent cover was quantified in FIJI by a researcher blind to burrow identity. Last, we recorded daily temperature fluctuations of additional (i.e., not included in the RFID study) bur-
row interiors using HOBO® U23 Pro v2 data loggers (Onset Computer Co., Bourne, MA). We recorded for two weeks from both natural and artificial burrows, created by coring sand plugs of various lengths (e.g., 10, 40, 60, and 80cm) from the sides of dunes at a 25-degree downward angle.

**Statistics.** We tested for sex differences in nightly activity level using a generalized linear mixed-effects model (GLMM) with RFID ping count as the response variable, sex as a fixed effect, and night, burrow ID, and mouse ID as random effects. We tested for sex differences in the number of burrows visited per night using a GLMM with sex as a fixed effect and mouse ID as a random effect. All GLMMs were run with Poisson error distributions and log link functions. We also tested for sex differences in the number of burrows visited in total using an LMM with sex as a fixed effect and the number of nights the mouse was detected as a random effect.

### C.2 Materials and Methods for May 2016 Fieldwork

#### Phase I: Capture-recapture (April 30th – May 10th 2016)

We set a 60 x 7 trap grid, with traps spaced at 15m intervals along both east-west and north-south axes. After the recommendation of Bondrup-Neilsen, we used a trap grid 16 times larger than the average home-range size. The final grid measured 885 x 90m (79,650 m²) and contained 420 trap locations. We trapped for 10 consecutive nights, but not all traps were set on all nights; see Table C4 for usage information. Altogether, we trapped 40 mice, 39 of which we successfully tagged and released.

#### Phase II: RFID monitoring (May 11th – May 18th 2016)

We recorded activity at a total of 29 unique burrows. Not all burrows were active on all nights; see Table C5 for details. We observed RFID activity for 79% (31/39) of the mice that were tagged and
Table C.4: May 2016 capture-recapture summary.

<table>
<thead>
<tr>
<th>Night</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traps set</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3990</td>
</tr>
<tr>
<td>Captures</td>
<td>210</td>
<td>420</td>
<td>420</td>
<td>420</td>
<td>420</td>
<td>420</td>
<td>420</td>
<td>420</td>
<td>420</td>
<td>420</td>
<td></td>
</tr>
<tr>
<td>New captures</td>
<td>3</td>
<td>6</td>
<td>2</td>
<td>7</td>
<td>12</td>
<td>12</td>
<td>16</td>
<td>16</td>
<td>14</td>
<td>16</td>
<td>112</td>
</tr>
<tr>
<td>Losses</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>40</td>
</tr>
<tr>
<td>Cumulative</td>
<td>3</td>
<td>8</td>
<td>10</td>
<td>13</td>
<td>18</td>
<td>19</td>
<td>22</td>
<td>26</td>
<td>32</td>
<td>40</td>
<td>40</td>
</tr>
</tbody>
</table>

Table C.5: May 2016 RFID summary.

<table>
<thead>
<tr>
<th>Night</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active RFID readers</td>
<td>10</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>20</td>
<td>16</td>
<td>17</td>
<td>29</td>
</tr>
<tr>
<td>No. mice detected</td>
<td>16</td>
<td>19</td>
<td>18</td>
<td>18</td>
<td>22</td>
<td>19</td>
<td>21</td>
<td>31</td>
</tr>
<tr>
<td>Total activity (no. pings)</td>
<td>190</td>
<td>599</td>
<td>929</td>
<td>222</td>
<td>364</td>
<td>489</td>
<td>237</td>
<td>3030</td>
</tr>
</tbody>
</table>

released in Phase I.

SECR model. To estimate home-range size and population density, we ran a spatially explicit capture-recapture analysis using the secr package in R. We included all captures from 10 consecutive trap nights. We used the “suggest.buffer” function to determine a suitable buffer width (116m) around the 885 x 90m trap grid. We then fit a half-normal spatial detection model by maximum likelihood using the “secr.fit” function, which estimates two parameters, $g_0$ and $\sigma$ (Table C6). We estimated the 95% home-range radius from $\sigma$ using the “circular.r” function.

Table C.6: May 2016 spatially explicit capture-recapture (SECR) model results.

<table>
<thead>
<tr>
<th>Value</th>
<th>estimate</th>
<th>s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prob. of capture at activity centre ($g_0$)</td>
<td>0.029</td>
<td>0.005</td>
</tr>
<tr>
<td>Scale parameter (s)</td>
<td>44.33</td>
<td>2.95</td>
</tr>
<tr>
<td>95% home-range radius (m)</td>
<td>99.94</td>
<td>-</td>
</tr>
<tr>
<td>Home-range size (m²)</td>
<td>31,375</td>
<td>-</td>
</tr>
<tr>
<td>Density (mice/km²)</td>
<td>157.6</td>
<td>26.9</td>
</tr>
</tbody>
</table>
**Burrow visitation patterns.** We limited the following analyses to mice with 4 or more pings and burrows with 6 or more pings over the entirety of the 7-night study. In total, we analysed data from 27 mice and 23 burrows. To identify which burrows showed similar visitation patterns, we first determined the portion of each burrow's total RFID activity comprised by different mice. We then scaled and centered these data and calculated a dissimilarity matrix based on Euclidean distance. We then used agglomerative hierarchical clustering to produce a cluster dendrogram of burrows visited by similar groups of mice. We next defined 9 burrow clusters by drawing a horizontal line through the dendrogram at a minimum height such that all burrows within a cluster shared at least one mouse. Mice frequented burrows belonging to 1-3 clusters (mean = 1.48).

### C.3 Results for May 2016 Fieldwork

**Population density and home-range size**

Using a spatially explicit capture-recapture model, we estimated a population density of 158 ± 27 mice/km². The probability of trapping a mouse at its home-range centre was 0.029 ± 0.005 and the mean home-range size estimate was 31,735 m² (Fig. C2c).

**Burrow visitation profiles**

**Total activity summaries.** We found no significant difference in the total number of burrows visited by males and females (LMM, P = 0.541). Mice visited and between 1 and 5 burrows over the course of the 7-night study (mean = 2.6 ± 0.2 s.e. burrows per mouse, n = 27 mice) (Fig. C3c) and we detected a total of 1 to 6 mice at each burrow (mean = 3.0 ± 0.3 s.e. mice per burrow, n = 23 burrows) (Fig. C3d).

**Nightly activity summaries.** We found no significant difference in nightly activity level
Figure C.2: Home-range characterization and burrow use patterns as determined by capture-recapture and RFID-monitoring. a, Satellite image depicting a section of Santa Rosa Island, FL (30°23'42" N, 86°43'27" W). The dotted line denotes the trap grid boundary (expanded in b). b, Trap grid captures and RFID-monitored burrow locations. Trap area polygons are drawn for 13/31 individuals that were captured at three or more unique trap locations (n = 8 females, n = 5 males). Grey circles denote RFID burrow locations (n = 29). Circle size is log2 scaled by the total number of RFID pings recorded over 7 nights. c, Half-normal detection function describing the capture probability as a function of distance from the home-range centre. Grey dashed line represents the standard error of the model estimate. Red dashed line represents the scale parameter value (σ). g0 = intercept, σ = scale parameter, hr = 95% home-range radius. Data from May 2016.

(i.e., number of RFID pings) between males and females (GLMM, P = 0.799). We also found no significant sex difference in the number of burrows visited per night (GLMM, P = 0.432). Mice visited between 1 and 4 burrows per night (mean = 1.5 ± 0.1 s.e. burrows per mouse, n = 27 mice) (Fig. C.4e) and, each night, we detected between 1 and 4 mice at each burrow (mean = 1.9 ± 0.1 s.e. mice per burrow, n = 23 burrows) (Fig. C.4f).
Burrow preferences. Over the course of the study, the majority of mice (23/27) were detected at more than one burrow. The mean distance between burrow sites in an individual’s burrow network was $84.3 \pm 17.0$ m ($n = 23$ mice) (Fig. C4c). However, not all burrows were used equally. On average, $72 \pm 4\%$ of a mouse’s total RFID activity was observed at a single burrow ($n = 27$ mice) (Fig. C4d). Thus, while mice consistently visited several burrows within their home-range—traveling up to 435 m between burrow sites—the majority of their activity was often restricted to a single burrow.

Figure C.3: Groups of mice visit networks of spatially clustered burrows. a, Dendrogram depicting hierarchical clustering of burrows visited by similar groups of beach mice. Burrow clusters are indicated by coloured circles placed at nodes. Each pie chart represents a unique burrow; pie size is log2 scaled by the total number of RFID pings recorded over 7 nights. Pie slices represent the portion of the total RFID burrow activity comprised by different mice (females in red/orange; males in blue/green). b, Spatial locations of burrow clusters depicted in a. Circle size is log2 scaled by the total number of RFID pings. c, Total number of burrows visited by each mouse over 7 nights ($n = 27$ mice). d, Total number of mice detected at each burrow over 7 nights ($n = 23$ burrows). Data from May 2016.
Figure C.4: Nightly activity summaries, distance between burrows, and burrow use evenness. a, Mean number of burrows visited per night for each mouse (n = 27 mice). b, Mean number of mice detected per night for each burrow (n = 23 burrows). c, Mean distance between all burrows each mouse's network (n = 27 mice). d, Portion of each mouse's total RFID activity observed at each burrow in its network (n = 27 mice). Data from May 2016.

Figure C.5: Distribution of RFID timestamps and fraction of spatiotemporal overlap in the population as a function of time-bin size. a, Histogram of RFID timestamps for female and male mice over the 7-night study. At the midpoint of the study, sunset and sunrise occurred at 7:33 PM and 5:52 AM, respectively. b, Network load (fraction of all possible pairs of mice that are detected at the same burrow entrance within a given time window). The green dot at 1.5 hours denotes the elbow of the curve and, accordingly, the time-bin size used to calculate overlap index for all downstream analyses. Data from May 2016.

Spatiotemporal overlap

We found evidence of spatiotemporal overlap for 20% (70/351) of possible pairs. The maximum OI was 0.036 and the mean OI (excluding zeroes) was 0.0087 (Fig. C6a). We found evidence of spatiotemporal overlap for 100% (351/351) of pairs in the randomized dataset. However, the magnitude of overlap was much lower than in the real dataset. The maximum OI was 0.012 and the mean OI
(excluding zeroes) was 0.0068 (Fig. C6b).

Figure C.6: Spatiotemporal overlap at burrow entrances. a, Heatmap of overlap indices (OIs) for all possible pairs of mice in the RFID-tracked population. The order of focal and partner mice is determined by hierarchical clustering. Females are indicated in red ($n = 12$) and males in blue ($n = 15$). Pairs of mice that were not observed at the same burrow within a 1.5-hour time window are indicated in navy (OI = 0), whereas pairs that overlapped extensively are shown in green/yellow. Legend at right applies to both a and b. The distribution of non-zero overlap indices is indicated below the heatmap. A dashed white line is drawn at the mean OI. b, Same data as in a, but with randomized mouse and burrow IDs assigned to RFID timestamps. Data from May 2016.
SOCIAL NETWORK ANALYSIS AND GENETIC RELATEDNESS

Pairs that overlapped (OI > 0) had higher coefficients of relatedness than those that did not (OI = 0) (Linear model, F = 4.43, P = 0.036) (Fig. C7b).

Figure C.7: Close relatives show greater spatiotemporal overlap. a, Overlap index (OI) network for the RFID-tracked population. Network nodes represent individual mice, ordered by hierarchical clustering of OI values. Females are indicated in red (n = 12) and males in blue (n = 15). Network edges represent non-zero OIs and are weighted according to OI value. b, Genetic relatedness network. Network nodes represent individual mice and are ordered as in a. Network edges are weighted according to the coefficient of relatedness. c, Relatedness coefficients for pairs of mice that did not overlap (white, n = 488) versus those that did (grey, n = 64). Significance levels: *P < 0.05. Error bars represent SEM. Data from May 2016.
Supplementary Material for Chapter 5
Figure D.1: Pearson correlation coefficients for pairs of component burrowing traits. Correlation matrices are same as in Figure 5.2. *P. maniculatus* (left); *P. polionotus* (right). Component traits are ordered according to hierarchical clustering of *P. polionotus* data based on Euclidean distance (dendrogram shown at top-right).
<table>
<thead>
<tr>
<th>Video traits</th>
<th>Category</th>
<th>Photo traits</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>burrow_area_excavated</td>
<td>BA</td>
<td>X1_area</td>
<td>BA</td>
</tr>
<tr>
<td>burrow_area_total</td>
<td>BA</td>
<td>X1_branch_count</td>
<td>BA</td>
</tr>
<tr>
<td>burrow_length_max</td>
<td>BA</td>
<td>X1_branch_length</td>
<td>BA</td>
</tr>
<tr>
<td>burrow_length_total</td>
<td>BA</td>
<td>X1_entrance_angle</td>
<td>BA</td>
</tr>
<tr>
<td>burrow_main_initiated</td>
<td>BA</td>
<td>X1_entrance_exit_difference</td>
<td>BA</td>
</tr>
<tr>
<td>burrow_main_peak_activity</td>
<td>BA</td>
<td>X1_exit_angle</td>
<td>BA</td>
</tr>
<tr>
<td>ground_accrued</td>
<td>BA</td>
<td>X1_exit_count</td>
<td>BA</td>
</tr>
<tr>
<td>ground_removed</td>
<td>BA</td>
<td>X1_fraction_upwards</td>
<td>BA</td>
</tr>
<tr>
<td>predug_area</td>
<td>BA</td>
<td>X1_goodBur</td>
<td>BA</td>
</tr>
<tr>
<td>time_spent_digging</td>
<td>BA</td>
<td>X1_length</td>
<td>BA</td>
</tr>
<tr>
<td>mouse_deepest_diagonal</td>
<td>MB</td>
<td>X1_length_upwards</td>
<td>BA</td>
</tr>
<tr>
<td>mouse_deepest_vertical</td>
<td>MB</td>
<td>X1_openness</td>
<td>BA</td>
</tr>
<tr>
<td>mouse_digging_rate</td>
<td>MB</td>
<td>X1_perimeter</td>
<td>BA</td>
</tr>
<tr>
<td>mouse_distance_covered</td>
<td>MB</td>
<td>X1_perimeter_exit</td>
<td>BA</td>
</tr>
<tr>
<td>mouse_running_peak</td>
<td>MB</td>
<td>X1_pos_x</td>
<td>BA</td>
</tr>
<tr>
<td>mouse_speed_max</td>
<td>MB</td>
<td>X1_pos_y</td>
<td>BA</td>
</tr>
<tr>
<td>mouse_speed_max_aboveground</td>
<td>MB</td>
<td>X1_total_length</td>
<td>BA</td>
</tr>
<tr>
<td>mouse_speed_max_underground</td>
<td>MB</td>
<td>X1X2_area</td>
<td>BA</td>
</tr>
<tr>
<td>mouse_speed_mean</td>
<td>MB</td>
<td>X1X2_length</td>
<td>BA</td>
</tr>
<tr>
<td>mouse_speed_mean_aboveground</td>
<td>MB</td>
<td>X1X2_total_length</td>
<td>BA</td>
</tr>
<tr>
<td>mouse_speed_mean_underground</td>
<td>MB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mouse_speed_mean_valid</td>
<td>MB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mouse_speed_mean_valid_aboveground</td>
<td>MB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mouse_speed_mean_valid_underground</td>
<td>MB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mouse_trail_longest</td>
<td>MB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>time_at_burrow_end</td>
<td>MB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>time_spent_aboveground</td>
<td>MB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>time_spent_moving</td>
<td>MB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>time_spent_moving_aboveground</td>
<td>MB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>time_spent_moving_underground</td>
<td>MB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>time_spent_underground</td>
<td>MB</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table D.2: QTL mapping summary for mouse behavior traits. Chr = Chromosome; LOD = Logarithm of the odds, \( P = P \) value, PVE = Percent variance explained; Peak = position (in base pairs) of the genetic marker with the highest LOD score; BCI = Bayes credible interval; Trans = transgressive QTL (true or false). Transgressive = NA if the component trait is not significantly different between the pure species.

<table>
<thead>
<tr>
<th>Component trait</th>
<th>Chr</th>
<th>LOD</th>
<th>( P )</th>
<th>PVE</th>
<th>Peak (bp)</th>
<th>BCI (Mb)</th>
<th>Trans</th>
</tr>
</thead>
<tbody>
<tr>
<td>mouse_running_peak</td>
<td>2</td>
<td>4.41</td>
<td>0.0383</td>
<td>3.56</td>
<td>149943572</td>
<td>87.1</td>
<td>F</td>
</tr>
<tr>
<td>mouse_speed_max</td>
<td>2</td>
<td>4.34</td>
<td>0.034</td>
<td>3.5</td>
<td>119519930</td>
<td>53.7</td>
<td>NA</td>
</tr>
<tr>
<td>mouse_speed_max_aboveground</td>
<td>2</td>
<td>4.34</td>
<td>0.034</td>
<td>3.5</td>
<td>119519930</td>
<td>53.7</td>
<td>NA</td>
</tr>
<tr>
<td>mouse_distance_covered</td>
<td>2</td>
<td>13.02</td>
<td>&lt;0.001</td>
<td>10.14</td>
<td>136427060</td>
<td>13.9</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>7.21</td>
<td>&lt;0.001</td>
<td>5.75</td>
<td>72969997</td>
<td>174.4</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>21</td>
<td>5.41</td>
<td>0.00319</td>
<td>4.34</td>
<td>45783430</td>
<td>77.9</td>
<td>NA</td>
</tr>
<tr>
<td>mouse_speed_mean</td>
<td>2</td>
<td>11.64</td>
<td>&lt;0.001</td>
<td>8.72</td>
<td>137306688</td>
<td>14.1</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>5.57</td>
<td>0.00319</td>
<td>4.47</td>
<td>72932893</td>
<td>178.4</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>21</td>
<td>5.53</td>
<td>0.00426</td>
<td>4.44</td>
<td>45783430</td>
<td>50.5</td>
<td>NA</td>
</tr>
<tr>
<td>time_spent_moving</td>
<td>2</td>
<td>11.11</td>
<td>&lt;0.001</td>
<td>9.11</td>
<td>136196821</td>
<td>14.1</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>6.98</td>
<td>&lt;0.001</td>
<td>5.57</td>
<td>72969997</td>
<td>134.3</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>21</td>
<td>5.89</td>
<td>0.00106</td>
<td>4.72</td>
<td>45783430</td>
<td>69.5</td>
<td>NA</td>
</tr>
<tr>
<td>mouse_speed_mean_aboveground</td>
<td>2</td>
<td>12.8</td>
<td>&lt;0.001</td>
<td>9.97</td>
<td>136342093</td>
<td>10.3</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>6.68</td>
<td>&lt;0.001</td>
<td>5.34</td>
<td>72932893</td>
<td>178</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>21</td>
<td>5.05</td>
<td>0.00106</td>
<td>4.06</td>
<td>45783430</td>
<td>92.4</td>
<td>NA</td>
</tr>
<tr>
<td>time_spent_moving_aboveground</td>
<td>2</td>
<td>11.68</td>
<td>&lt;0.001</td>
<td>9.14</td>
<td>137306688</td>
<td>15.1</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>6.68</td>
<td>&lt;0.001</td>
<td>5.34</td>
<td>72969997</td>
<td>169.4</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>21</td>
<td>6.46</td>
<td>&lt;0.001</td>
<td>5.16</td>
<td>45783430</td>
<td>50.6</td>
<td>NA</td>
</tr>
<tr>
<td>time_spent_aboveground</td>
<td>21</td>
<td>4.62</td>
<td>0.0202</td>
<td>3.72</td>
<td>46487</td>
<td>46.4</td>
<td>T</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>6.79</td>
<td>&lt;0.001</td>
<td>5.42</td>
<td>77533814</td>
<td>75.5</td>
<td>T</td>
</tr>
<tr>
<td>mouse_speed_max_underground</td>
<td>1</td>
<td>4.87</td>
<td>0.00958</td>
<td>3.92</td>
<td>132710587</td>
<td>98.5</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>6.24</td>
<td>&lt;0.001</td>
<td>4.99</td>
<td>5511143</td>
<td>79.2</td>
<td>T</td>
</tr>
<tr>
<td>time_spent_moving_underground</td>
<td>1</td>
<td>5.18</td>
<td>0.00745</td>
<td>4.16</td>
<td>132839023</td>
<td>52.6</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>9.32</td>
<td>&lt;0.001</td>
<td>7.37</td>
<td>70708951</td>
<td>41.9</td>
<td>T</td>
</tr>
<tr>
<td>time_spent_underground</td>
<td>21</td>
<td>7.03</td>
<td>&lt;0.001</td>
<td>5.61</td>
<td>25093694</td>
<td>51.2</td>
<td>T</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>9.2</td>
<td>&lt;0.001</td>
<td>7.27</td>
<td>77642472</td>
<td>46.2</td>
<td>T</td>
</tr>
<tr>
<td>time_spent_digging</td>
<td>1</td>
<td>4.31</td>
<td>0.0308</td>
<td>3.48</td>
<td>132924018</td>
<td>80</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td>21</td>
<td>6.51</td>
<td>&lt;0.001</td>
<td>5.2</td>
<td>25093694</td>
<td>50.9</td>
<td>T</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>8.79</td>
<td>&lt;0.001</td>
<td>6.96</td>
<td>77642472</td>
<td>34</td>
<td>T</td>
</tr>
<tr>
<td>time_at_burrow_end</td>
<td>21</td>
<td>6.2</td>
<td>&lt;0.001</td>
<td>4.96</td>
<td>23905115</td>
<td>13.8</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>11.2</td>
<td>&lt;0.001</td>
<td>8.78</td>
<td>77601976</td>
<td>51.3</td>
<td>NA</td>
</tr>
<tr>
<td>burrow_main_peak_activity</td>
<td>X</td>
<td>4.36</td>
<td>&lt;0.001</td>
<td>3.52</td>
<td>49204476</td>
<td>29.5</td>
<td>T</td>
</tr>
</tbody>
</table>
Table D.3: QTL mapping summary for burrow architecture traits. Chr = Chromosome; LOD = Logarithm of the odds, 
P = P value, PVE = Percent variance explained; Peak = position (in base pairs) of the genetic marker with the highest 
LOD score; BCI = Bayes credible interval; Trans = transgressive QTL (true or false). Transgressive = NA if the compo-
nent trait is not significantly different between the pure species.

<table>
<thead>
<tr>
<th>Component trait</th>
<th>Chr</th>
<th>LOD</th>
<th>P</th>
<th>PVE</th>
<th>Peak (bp)</th>
<th>BCI (Mb)</th>
<th>Trans</th>
</tr>
</thead>
<tbody>
<tr>
<td>mouse_deepest_diagonal</td>
<td>1</td>
<td>4.33</td>
<td>0.0393</td>
<td>3.49</td>
<td>61161124</td>
<td>97.4</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>11.03</td>
<td>&lt;0.001</td>
<td>8.66</td>
<td>72067712</td>
<td>26.9</td>
<td>T</td>
</tr>
<tr>
<td>mouse_deepest_vertical</td>
<td></td>
<td>4.75</td>
<td>0.0149</td>
<td>3.82</td>
<td>65812562</td>
<td>83.4</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>14.09</td>
<td>&lt;0.001</td>
<td>10.92</td>
<td>72127458</td>
<td>18.1</td>
<td>T</td>
</tr>
<tr>
<td>mouse_trail_longest</td>
<td>X</td>
<td>8.04</td>
<td>&lt;0.001</td>
<td>6.39</td>
<td>73172274</td>
<td>44.9</td>
<td>T</td>
</tr>
<tr>
<td>burrow_length_total</td>
<td>X</td>
<td>9.74</td>
<td>&lt;0.001</td>
<td>7.68</td>
<td>79044466</td>
<td>34</td>
<td>T</td>
</tr>
<tr>
<td>burrow_length_max</td>
<td>X</td>
<td>10.1</td>
<td>&lt;0.001</td>
<td>7.96</td>
<td>79047461</td>
<td>44.8</td>
<td>T</td>
</tr>
<tr>
<td>burrow_area_total</td>
<td>X</td>
<td>7.9</td>
<td>&lt;0.001</td>
<td>6.28</td>
<td>79044466</td>
<td>26.4</td>
<td>T</td>
</tr>
<tr>
<td>burrow_area_excavated</td>
<td>X</td>
<td>8.15</td>
<td>&lt;0.001</td>
<td>6.47</td>
<td>79044466</td>
<td>31.8</td>
<td>T</td>
</tr>
<tr>
<td>Xi_openness</td>
<td></td>
<td>4.37</td>
<td>0.034</td>
<td>3.52</td>
<td>37416298</td>
<td>159.3</td>
<td>F</td>
</tr>
<tr>
<td>Xi_perimeter</td>
<td>X</td>
<td>8.06</td>
<td>&lt;0.001</td>
<td>6.4</td>
<td>75031651</td>
<td>33.7</td>
<td>T</td>
</tr>
<tr>
<td>Xi_entrance_exit_difference</td>
<td>X</td>
<td>6.98</td>
<td>&lt;0.001</td>
<td>5.57</td>
<td>81980958</td>
<td>26.7</td>
<td>T</td>
</tr>
<tr>
<td>Xi_total_length</td>
<td>X</td>
<td>8.47</td>
<td>0.00319</td>
<td>3.6</td>
<td>81764195</td>
<td>33.9</td>
<td>T</td>
</tr>
<tr>
<td>Xi_length</td>
<td>X</td>
<td>8.72</td>
<td>&lt;0.001</td>
<td>6.91</td>
<td>77533814</td>
<td>24.5</td>
<td>T</td>
</tr>
<tr>
<td>Xi_area</td>
<td>X</td>
<td>8.42</td>
<td>&lt;0.001</td>
<td>6.68</td>
<td>77533814</td>
<td>22.6</td>
<td>T</td>
</tr>
<tr>
<td>Xi_area</td>
<td>X</td>
<td>4.89</td>
<td>0.00213</td>
<td>3.93</td>
<td>79933807</td>
<td>40.1</td>
<td>T</td>
</tr>
</tbody>
</table>