Collective sperm movements are shaped by post-copulatory sexual selection and phylogenetic history in *Peromyscus* mice

Kristin A. Hook, W. David Weber, Heidi S. Fisher*

Department of Biology, University of Maryland, 1200 Biology-Psychology Building, 4094 Campus Drive, College Park, MD 20742, U.S.A.

* Heidi S. Fisher
Email: hsfisher@umd.edu

Keywords:

mating systems; sexual selection; sperm competition; sperm conjugation; sperm motility

Author Contributions:

KAH and HSF conceived of the study, designed experiments, and interpreted results; KAH and WWD collected the data, KAH carried out the statistical analyses; all authors wrote the manuscript, gave final approval for publication, and agree to be held accountable for the work presented.

Funding:

This work was supported by a Eunice Kennedy Shriver National Institute of Child Health and Human Development K99/R00 Pathway to Independence Award to HSF [R00HD071972] and a National Science Foundation Postdoctoral Research Fellowship to KAH [1711817].

Abstract

Sperm of some species form motile, coordinated groups as they migrate through the female reproductive tract to the site of fertilization. This collective motion is predicted to improve sperm swimming performance and therefore may be beneficial in a competitive context, but limited evidence supports this theory. Here we examine sperm aggregates across closely-related species of *Peromyscus* mice that naturally vary by mating system, and thus sperm competition intensity. We find that phylogenetic history predicts the likelihood that sperm will aggregate, and that relative testis size is negatively associated with variation in number of aggregated cells, suggesting that sperm competition has a stabilizing effect on sperm group size. Moreover, we show that aggregates are kinematically beneficial for some species but costly for others, and these differences are largely dependent on the orientation and composition of sperm within the groups. In addition, when we compared sperm of the two sister-species that aggregate most frequently, we find that sperm from the species that evolved under intense sperm competition forms aggregates with more efficient geometry more frequently than sperm from its monogamous congener. These results are consistent with the prediction that sperm aggregation evolved to improve motility in a competitive context; however, when monogamy evolved secondarily, relaxed sexual selection allowed for less efficient strategies to persist. Together, our findings in *Peromyscus* reveal that collective sperm behavior is likely to evolve rapidly and is shaped by changes in the selective regime.

Introduction

Sperm cells are one of the most diverse cell types in nature and exhibit striking variation both within and across taxa (1). In addition to being morphologically diverse, sperm may exhibit complex, emergent behaviors, including sperm conjugation, which occurs when two or more cells join together for motility or transport through the female reproductive tract before dissociating prior to fertilization (1, 2). Although relatively rare, these sperm-sperm interactions have evolved

74 multiple times across independent lineages of internally fertilizing species (3, 1, 2), yet the
75 adaptive significance of these gametic interactions remains unclear for many taxa.
76 One functional hypothesis posits that sperm aggregates may be advantageous if the combined
77 force generated by multiple flagella enable sperm conjugates to swim faster than single cells (4, 5).
78 Improved sperm motility can be beneficial if it enables cells to move quickly through hostile
79 environments within the female tract (6, 7), and in competitive environments in which females
80 mate with multiple males, as sperm velocity is often correlated with competitive fertilization
81 success (e.g., 7–9). While some studies that have quantified the motility of sperm aggregates have
82 found empirical support for this hypothesis (e.g., 4, 10–13), others have not (e.g., 14, 15). For
83 example, sperm groups swim faster than solitary cells in the Norway rat (*Rattus norvegicus*), yet
84 house mouse (*Mus musculus*) sperm swim slower as groups under identical experimental
85 conditions (5). Another non-mutually exclusive hypothesis is that sperm aggregation facilitates
86 migration through viscous or viscoelastic secretions of the female reproductive tract (16, 17),
87 including cervical mucus (18), which has received some empirical support (e.g., 4, 12, 16), but see
88 (15). In the grey short-tailed opossum (*Monodelphis domestica*), for example, sperm pairs swim
89 with greater motility than single sperm in viscous fluids; however, these sperm pairs were
90 artificially induced to uncouple (16) and therefore the comparison may not be biologically
91 informative (1). Additionally, sperm aggregation may protect sensitive regions of the sperm, such
92 as the acrosome, from damage and preserve sperm functionality during passage through the male
93 or female reproductive tracts (19–22), or enhance egg penetration during fertilization (23) but see
94 (24). While it is often assumed that coordinated sperm movements are adaptive, inconsistent
95 findings and a wide diversity of naturally variable sperm behavior have limited our understanding
96 of the functional advantages of sperm aggregation.
97 Multiple independent origins of sperm conjugation suggest that the functional consequences,
98 as well as mechanisms that regulate these cellular interactions, are likely to vary throughout nature
99 (reviewed in 2). Indeed, the formation of sperm aggregates and the number of grouped cells varies

100 widely across taxa (7). In mammals alone, sperm of some species assemble during epididymal
101 transport and are molecularly "glued" to one another as bi-flagellate pairs in grey short-tailed
102 opossums (25), as bundles of roughly 100 cells in monotremes (26), or as organized rouleaux of
103 five or more cells in guinea pigs (*Cavia porcellus*, 24). Conversely, mammalian sperm may also
104 assemble after ejaculation to form variably-sized groups. For instance, sperm may form temporary
105 clusters of up to sixteen cells in bulls (*Bos Taurus*, 15) or more fixed groups of up to 30 cells in
106 house mice (5), 50 cells in the Norway rat (5), or thousands of cells in the wood mouse (*Apodemus*
107 *sylvaticus*, 4) whereby the hook-shaped heads interlock or attach to the flagella of other sperm. For
108 these latter species in which sperm groups form after ejaculation, both single and aggregated
109 sperm typically co-occur, thus allowing for direct comparisons between collective and solitary
110 sperm movements within an ejaculate while controlling for within-male variability.

111 Closely-related species in the rodent genus, *Peromyscus,* produce sperm that naturally vary in
112 their collective behavior. Sperm of some species assemble temporary groups after ejaculation by
113 adhering to one another at their head region (27) and disassemble prior to fertilization (10). In *P.*
114 *maniculatus*, sperm selectively group with the most closely-related cells to form motile groups of
115 up to 30 cells (10), but there is a non-monotonic association between group size and swimming
116 velocity, indicating that some groups are faster than single cells but others are not (28).
117 Conversely, sperm produced by their sister-species, *P. polionotus*, also form aggregates but do so
118 indiscriminately with related and unrelated cells (10), are less likely to be optimally-sized, and
119 overall move in a less linear trajectory (28). Intriguingly, sperm competition is predicted to be
120 more intense in *P. maniculatus* than in *P. polionotus* due to their different mating systems. In *P.*
121 *maniculatus*, both sexes mate with multiple partners, often in overlapping series just minutes apart
122 (29), and females frequently carry multiple-paternity litters in the wild (30), whereas both
123 behavioral (31) and genetic data (32) indicate that *P. polionotus* is strictly monogamous. Evidence
124 suggests that monogamy has evolved at least twice within the *Peromyscus* lineage (reviewed in 33;
125 34), thus enabling us to investigate if post-copulatory sexual selection, driven by female mating

126 behavior, has shaped the evolution of sperm aggregation (4) more broadly across the *Peromyscus*
127 lineage. In this study, we quantify intra- and inter-specific differences in the size and performance
128 of sperm aggregates under consistent, controlled conditions to examine the evolution of collective
129 sperm behavior and empirically test whether sperm aggregation improves swimming performance.
130

131 **Results**

132 We investigated the frequency of sperm aggregation in each species and found significant
133 differences in the proportion of aggregated cells among species (binomial GLMM: $n = 134$, $P <$
134 0.001; Table 1, 2; Figure 1), with more variance across species than within species (variance
135 across species = 0.10; variance within each species < 0.10). Specifically, sperm from *P.*
136 *maniculatus* and *P. polionotus* similarly aggregate the most, whereas *P. leucopus* and *P.*
137 *gossypinus* sperm similarly aggregate the least among all the species. Pairwise comparisons
138 adjusted for multiple comparisons using LSmeans revealed that species within these pairs do not
139 significantly differ from one another, whereas all other pairwise species comparisons do (Table 1,
140 2). The coefficient of variation (CV) for the proportion of cells aggregated within each species are
141 as follows: 68.8% for *P. californicus*, 50.6% for *P. eremicus*, 14.4% for *P. polionotus*, 10.8% for
142 *P. maniculatus*, 125.7% for *P. leucopus*, and 60.5% for *P. gossypinus*. Controlling for phylogenic
143 relationships and body mass, we found no effect of testis size ($F_{2,3} = 2.606$, $P = 0.15816$) or the
144 within-species CV ($F_{2,3} = 0.3603$, $P = 0.5604$) on the proportion of aggregated cells.
145 Moreover, we found significant differences in the mean number of cells aggregated among species
146 (LM: $F_{6,126} = 56.37$, $P < 0.001$; Table 1, 2; Figure 2A), with more variance observed across species
147 ($s^2 = 1.96$) than within species ($s^2 < 1.00$, except for *P. polionotus* [$s^2 = 2.15$]). Post-hoc pairwise
148 comparisons revealed that both *P. maniculatus* and *P. polionotus* produce the largest sperm
149 aggregates ($P < 0.05$ for all pairwise comparisons), whereas *P. gossypinus*, *P. leucopus*, and *P.*
150 *californicus* produce the smallest aggregates ($P < 0.05$ for pairwise comparisons), the latter of
151 which produces sperm cell aggregates that are statistically similar in size to those produced by *P.*

152 *eremicus* (P = 0.9203; Figure 2A). Controlling for phylogenic relationships and body mass, we
153 found a significant effect of testis weight on the within-species CV of aggregate size ($F_{2,3}$ = 8.398,
154 P = 0.02655, Figure 2B), but found less of an effect on the direct measure of aggregate size ($F_{2,3}$ =
155 3.058, P = 0.1522). The CV for the number of cells aggregated within each species were as
156 follows: 26.5% for *P. californicus*, 19.2% for *P. eremicus*, 27.5% for *P. polionotus*, 18.4% for *P.*
157 *maniculatus*, 13.3% for *P. leucopus*, and 7.2% for *P. gossypinus*.

158 By comparing sperm aggregate composition and orientation within the species that produce
159 the largest and most frequent sperm aggregates, we found that there are significantly more sperm
160 aggregates in *P. maniculatus* in which all sperm are aligned in a head-to-flagella orientation
161 (99.1%, 731/738) than in *P. polionotus* (87.5%, 720/826; GLM: n = 42, P = 3.92e-12) and that
162 these aggregates have significantly faster speeds (VCL) compared to aggregates with unaligned
163 cell orientations in both species (Figure 3; paired t-tests: *P. maniculatus* t = 5.9627, df = 4, P =
164 0.003972; *P. polionotus* t = 11.247, df = 11, P = 2.257e-07). We also found that there are
165 significantly fewer sperm aggregates in *P. maniculatus* with immotile, stuck, or morphologically
166 abnormal cells (8.1%, 64/795) than in *P. polionotus* (11.5%, 94/814; binomial GLM: n = 40, P =
167 0.00503). Importantly, aggregates with these defects had significantly lower speeds (VCL)
168 compared to aggregates without in both species (*P. maniculatus* $VCL_{aligned}$ = 179.60 ± 6.88 μm/sec,
169 $VCL_{defective}$ = 132.47 ± 6.18 μm/sec, paired t-test: t = 20.627, df = 14, P = 7.075e-12; *P. polionotus*
170 $VCL_{aligned}$ = 135.48 ± 3.51 μm/sec, $VCL_{defective}$ = 115.62 ± 5.22 μm/sec, paired t-test: t =16.312, df
171 = 20, P = 5.079e-13).

172 When comparing the proportion of motile and progressively motile aggregates across species,
173 our pairwise comparisons revealed that *P. eremicus* produced a significantly smaller proportion of
174 motile aggregates than all other species ($P < 0.05$ for all *P. eremicus* pairwise comparisons; $P >$
175 0.05 for all other pairwise comparisons). Fitted values of the proportion of motile aggregates using
176 LSmeans were 0.77 ± 0.04 for *P. eremicus*, 0.92 ± 0.02 for *P. polionotus*, 0.91 ± 0.02 for *P.*
177 *gossypinus*, 0.95 ± 0.01 for *P. maniculatus*, 0.93 ± 0.02 for *P. leucopus*, and 0.91 ± 0.02 for *P.*

178 *californicus*. Moreover, post-hoc comparisons revealed that *P. eremicus, P. polionotus,* and *P.*
179 *gossypinus* all had the smallest proportions of progressively motile aggregates ($P < 0.05$ for all
180 pairwise comparisons; fitted values using LSmeans were 0.64 ± 0.06, 0.65 ± 0.05, and 0.75 ± 0.05,
181 respectively), the latter species of which did not significantly differ from *P. californicus* ($P =$
182 0.1981; 0.87 ± 0.02 for LSmeans fitted values). Conversely, *P. maniculatus* and *P. leucopus* had
183 the largest proportions of progressively motile aggregates ($P = 0.3417$; fitted values using
184 LSmeans were 0.94 ± 0.02 and 0.91 ± 0.02, respectively), the latter of which did not differ from *P.*
185 *californicus* ($P = 0.9109$).

186 Overall, we found species-specific differences in the effect of sperm aggregation on motility,
187 regardless of environmental complexity (Table 3, Figure 4). In low-viscosity medium, we found
188 that sperm aggregates have a significantly greater VCL in *P. maniculatus*, VSL in *P. maniculatus*,
189 *P. leucopus*, and *P. californicus*, LIN in *P. leucopus*, and VAP in *P. californicus* compared to
190 single cells. Conversely, sperm aggregates had a significantly lower VCL, VSL, and VAP velocity
191 in *P. polionotus* and *P. gossypinus* than single sperm in the low-viscosity medium (Figure 4). In
192 the high-viscosity medium, we found that sperm aggregates have a significantly greater VCL in *P.*
193 *maniculatus* and *P. californicus* and a higher VSL and VAP in *P. californicus* (Figure 4) compared
194 to single sperm. Conversely, sperm aggregates in the high viscosity medium had a significantly
195 lower LIN in *P. californicus*, *P. eremicus*, *P. polionotus*, and *P. maniculatus* as well as a reduced
196 VSL and VAP in *P. polionotus* than single cells (Figure 4).

197

198 **Discussion**

199 While it is known that collective sperm behaviors have evolved independently in a number of
200 taxa (2), it remains unclear how sperm aggregation evolves among closely related species. Our
201 comparative study reveals that sperm aggregating behaviors vary across mice in the genus
202 *Peromyscus*. We observed an effect of phylogenetic history on the frequency of sperm
203 aggregation, indicating that collective sperm behavior likely evolved prior to the divergence of

present-day species. Additionally, we find a negative association between relative testis weight, a robust proxy for intensity of sperm competition in rodents (35), and the coefficient of variation for the number of cells in aggregate, suggesting that sexual selection has a stabilizing effect on sperm aggregate size. We then compared the motility of single sperm and sperm aggregates across all species under low- and high-viscosity conditions, and show that aggregation is kinematically beneficial for some species yet costly for others, regardless of environmental complexity. When examining sperm from the species that aggregate the most profoundly, we find that the formation of the aggregates and the orientation of the cells within the group are critical to kinematics. Moreover, we observe more aggregates with efficient geometry in the species that has evolved under strong post-copulatory sexual selection compared to its monogamous sister-species. These findings support the prediction that sperm aggregation evolved in *Peromyscus* to improve motility in a competitive context but reveal that relaxed selection may have enabled less efficient strategies to persist, thereby generating diversity in collective sperm behaviors within these closely-related species.

Our results reveal distinct species-specific differences across *Peromyscus* mice in the frequency of sperm aggregation and the average size of these cellular groups. Multiple *Peromyscus* species produce sperm that aggregate more extensively than other studied muroid rodents (36), with the notable exception of the wood mouse (4). In general, the proportion of sperm cells that aggregate is most similar within each sister-species pair. Specifically, *P. maniculatus* and *P. polionotus* produce sperm that aggregate the most (>80% of sperm), but the pair they are most closely-related to, *P. gossypinus* and *P. leucopus*, are the species least likely to produce sperm that aggregate (<10% of sperm); the most distantly-related species pair we assessed, *P. californicus* and *P. eremicus*, both produce sperm with a moderate propensity to aggregate (~25% of sperm). These findings support one of two possibilities for the evolution of sperm aggregates within these species: (a) a genus-wide ancestral trait of moderate sperm aggregation with subsequent diversification leading to an increase in *P. maniculatus* and *P.*

230 *polionotus* and a decrease in *P. gossypinus* and *P. leucopus*, or (b) the independent evolution of
231 aggregation in the ancestors of the *P. californicus*-species pair and the *P. maniculatus*-species pair.
232 Such complex evolutionary histories with losses and recurrences of sperm conjugation, and
233 subsequent species divergence, have also been demonstrated in diving beetles (Dytiscidae, 37),
234 consistent with the evolutionary lability that we observe. Our experimental results may explain
235 selection against sperm aggregation if forming groups reduces sperm swimming performance,
236 which we find in at least one species, *P. gossypinus*. However, we found that sperm produced by
237 their sister-species, *P. leucopus*, also rarely aggregate, despite our observation that these rare
238 collective groups have a greater average velocity and are more linear than single cells.
239 Interestingly, we observed that these two promiscuous species that rarely aggregate, *P. gossypinus*
240 and *P. leucopus*, have the largest relative testes of the species studied, suggesting that they may
241 have evolved increased sperm production to improve competitive fertilization success (sensu 38),
242 rather than adaptations that influence motility (39). Together these results suggest different
243 strategies employed by divergent species across the *Peromyscus* genus in response to sperm
244 competition.
245 In contrast to our results on the frequency of sperm aggregation within species, we found that
246 the average size of sperm groups does not align as closely with phylogenetic relationships. In
247 addition, we found that species with relatively larger testes, which is positively associated with
248 increased sperm competition (35), exhibit less variation (CV) in aggregate size. This result
249 supports the prediction that relaxed sperm competition allows for greater intermale variation to
250 persist in a population (40) and suggests that this post-copulatory sexual selection may be
251 stabilizing sperm aggregate size for a species-specific 'optimum' (28). Similarly, other studies
252 have shown that the strength of sexual selection regulates variance in sperm morphology across
253 taxa and at multiple levels of organization, including within- and between-males (41–43) as well
254 as within- and between-species (40, 44). A study on sperm bundles across ten *Carabus* ground
255 beetles also found intense selection on bundle size, which are dimorphic and either small or large;

256 the large, but not small, sperm bundles are positively correlated with measures of sperm
257 competition risk, including copulatory piece length and mate guarding, suggesting that diversity of
258 large sperm bundles is associated with sexual selection (45). In contrast to these findings that
259 competition drives sperm-sperm interactions, a study on the evolution of such sperm traits in
260 diving beetles found that variation in sperm conjugation is more associated with female
261 reproductive tract architecture (37, 46). Therefore, while our results suggest that stabilizing
262 selection on sperm aggregate size is associated with an increase in sperm competition given the
263 correlation with relative testis weight, mechanisms of female control (47) may also play an
264 important evolutionary role.

265 We compared the motility of single and aggregated sperm sampled from the same male to test
266 whether sperm aggregates swim faster or more efficiently than single sperm, which is predicted if
267 the combined force of multiple flagella enhances their motility (4). We found improved kinematic
268 measures in half of the species studied (*P. maniculatus, P. californicus*, and *P. leucopus*), thus
269 supporting this functional hypothesis; however, in several species we found aggregation had some
270 negative (*P. polionotus* and *P. gossypinus*) or no (*P. eremicus*) impact on motility relative to single
271 sperm cells. While we found support for a theoretical prediction that sperm aggregates achieve
272 greater straight-line velocity because they move in a more linear path of travel rather than at a
273 faster speed (i.e., curvilinear velocity, 28) in two of our six focal species (*P. californicus* and *P.*
274 *leucopus*), we did not find this kinematic benefit for sperm aggregation in all *Peromyscus* species.
275 These results corroborate other studies in more disparate taxonomic groups that have quantified
276 sperm aggregation motility and found inconsistent results. For example, sperm trains exhibit
277 greater swimming progressive motility in the wood mouse (4), and greater velocity than individual
278 sperm in the Norway rat, but not the house mouse (5). In invertebrates, the swimming velocity of
279 fishfly sperm increases with number of sperm in a bundle (12), but in a marine snail, there is no
280 differences in swimming speed between paired and single sperm (14). One possible explanation
281 for these differences across taxa is that cell orientation within an aggregate is critical for its

282 collective motility. Sperm cells are predicted to be faster if they generate increased force with
283 proportionally less drag (see 2 and references therein); such effects may be true for sperm
284 aggregations as well in which cells conjoin head-to-tail, thereby increasing the length of the
285 collective unit, or in which flagella within the group beat synchronously (2). Our results support
286 that even sperm aggregates that conjoin head-to-head and are thus wider, as they are in
287 *Peromyscus,* can offer a motility advantage.

288 Our results suggest that relaxed sexual selection may allow the persistence of less optimal
289 strategies based on the quantitative and qualitative differences we observed among sperm
290 aggregates of different species. The most interesting motility results are those of the sister-species
291 pair with divergent mating systems that both form the most frequent and largest sperm aggregates.
292 In *P. maniculatus*, a promiscuous species, sperm aggregates exhibit greater straight-line and
293 curvilinear velocity compared to single cells, but the opposite was true for its monogamous
294 congener, *P. polionotus*. We find that these kinematic differences are associated with differences
295 in aggregate geometry; when sperm heads and flagella are not oriented in the same direction, the
296 cells within an aggregate exert opposing forces on one another, thereby reducing the overall
297 motility of the group (27, 28). Indeed, we found that sperm from the monogamous *P. polionotus*
298 males are less likely to form aggregates with all sperm aligned and more likely to include immotile
299 or morphologically abnormal sperm, consequently resulting in slower aggregates than those of *P.*
300 *maniculatus* (27). This finding is consistent with previous reports that *P. polionotus* sperm tend to
301 form optimal-sized aggregates less often than in *P. maniculatus* (28). Together, these observations
302 further support the hypothesis that sperm aggregation evolved prior to the divergence of the
303 species pair (10), and when monogamy evolved secondarily in *P. polionotus* (34, 48), relaxed
304 sexual selection allowed for the persistence of less motile sperm traits. In line with this prediction,
305 we observed the smallest proportion of motile and progressively motile sperm aggregates in
306 another monogamous species, *P. eremicus,* but the largest proportion of progressively motile
307 aggregates in two promiscuous species, *P. maniculatus* and *P. leucopus*. Similar results have been

308 reported in house mice, in which males evolving under intense sperm competition produced a
309 greater proportion of motile sperm, compared to males from lineages subject to relaxed selection
310 (49). Together these findings support that the motility benefits conferred by sperm aggregation are
311 associated with variation in post-copulatory sexual selection.

312 When we compared single and aggregated sperm in a viscous environment to test if
313 aggregation improves motility in more complex fluids, similar to the secretions or cervical mucus
314 (18) of the female reproductive tract (17, 50), we found that collective motion is beneficial for
315 some species but costly for others. Compared to single sperm, aggregates swim with greater
316 curvilinear velocity in *P. maniculatus* and *P. californicus*, and greater straight-line and average
317 path velocities in *P. californicus,* but aggregates were less linear than single cells in all species
318 except for *P. leucopus* and *P. gossypinus*. While other studies have found kinematic benefits for
319 sperm conjugates compared to single cells in higher viscosities in the gray short-tailed opossum
320 (16), the wood mouse (4), and the fishfly (12), a study in bulls found that sperm were slower,
321 exhibited less organized swimming patterns, and were less likely to cluster in viscous fluids (15),
322 both of which are consistent with our analysis of *Peromyscus* aggregates. Ultimately, the benefit of
323 sperm aggregation depends on the relative importance of each kinematic parameter during sperm
324 migration *in vivo*. Although beyond the scope of this study, we predict that improved linearity
325 afforded by collective motion may help to direct the sperm through the dynamic fluids of the
326 female reproductive tract (51) and that increased velocity will reduce the time it takes for the
327 sperm to arrive at the fertilization site.

328 In conclusion, our study highlights the diversity of sperm aggregation within a single
329 taxonomic lineage and how selection has shaped the formation and performance of these cellular
330 groups. We show that both evolutionary history and varying levels of post-copulatory sexual
331 selection influence the frequency and size of sperm groups. Moreover, we find that sperm
332 aggregation can improve sperm motility in both simple and complex fluids, but this is not
333 consistent across all species. Theoretical predictions (27, 28) and emerging empirical evidence

334 suggests that motility benefits may only be realized if cells maintain optimal alignment within the
335 groups and, if achieved, may provide these sperm with a competitive advantage in the female
336 reproductive tract (46). Future work investigating sperm aggregates *in vivo* (e.g., 51, 52) will shed
337 light on the co-evolution of these unique gametic behaviors and the enormously variable and
338 dynamic female reproductive tracts through which sperm must navigate.
339

340 **Materials and Methods**

341 (a) Sperm collection

342 We obtained captive *Peromyscus maniculatus bairdii, P. polionotus subgriseus, P. leucopus*,
343 *P. eremicus*, and *P. californicus insignis* from the Peromyscus Genetic Stock Center at the
344 University of South Carolina, and *P. gossypinus gossypinus* from Dr. Hopi Hoekstra at Harvard
345 University and housed them in same-sex cages at 22ºC on a 16L:8D cycle in accordance with
346 guidelines established by the Institutional Animal Care and Use Committee at the University of
347 Maryland in College Park (protocol R-Jul-18-38). We sought samples from all available captive
348 *Peromyscus* species and avoided wild-caught specimens to control for variation due to life
349 experience. We obtained sperm samples from sexually mature males and accounted for relatedness
350 among the focal males by assigning siblings a unique 'Family' ID. We euthanized males via
351 isoflurane overdose and cervical dislocation, then weighed each male and both testes (Mettler
352 Toledo, Switzerland). Next, we removed a single caudal epididymis, made several small incisions
353 in the tissue, and submersed it in sperm medium (Modified Sperm Washing Medium, Irvine
354 Scientific, USA) that was pre-warmed at 37ºC; to reduce differences in sperm density despite
355 natural variation in epididymal sizes, we varied the volume (50µl - 1000µl) based on tissue size
356 and accounted for these differences when estimating final sperm counts for each male. To collect
357 sperm, we agitated the tissue at 300rpm (ThermoMixer F1.5, Eppendorf, Germany) at 37ºC for ten
358 minutes, inverting the tube at the five- and ten-minute mark, then incubated the tissue undisturbed
359 for two minutes. Using pipette tips cut to create a wider opening, we collected live sperm cells for

360 analysis from just below the meniscus of the solution to enrich for the most motile sperm (53).
361 Next we estimated sperm density using a computer-assisted sperm analysis (CASA) system (Ceros
362 II Animal, Hamilton Thorne, USA) and verified with a Neubauer-improved hemocytometer
363 (Marienfeld, Germany) then diluted samples with pre-warmed medium to reach a standard
364 concentration of 300-400 cells summed across the five 5-second videos at 100X magnification for
365 cell tracking optimality and efficiency.
366
367 (b) Live sperm analysis
368 To conduct live sperm observations, we gently reverse pipetted 4µl of the sperm solution into
369 12µl of pre-warmed medium on a plastic slide within a 9mm x 0.12mm imaging spacer (Grace
370 Bio-Labs, USA) and covered by a plastic cover slip. This set-up served as a control and represents
371 a 'low-viscosity' environment. To test sperm motility in a 'high-viscosity' environment, we
372 followed the same procedures except that we mixed 4µl of sperm solution with 12µl of pre-
373 warmed medium enriched with methylcellulose (Sigma Aldrich M 7140; 15cP, 2% in water; 54).
374 We then recorded 5-second videos at 60 frames/sec on the CASA system, capturing at least five
375 videos per male but recorded additional videos for samples with lower sperm density (n = 57) to
376 ensure an adequate number of observed cells per male. Videos (see example, Movie S1) were
377 recorded at 59±16 minutes post-harvest from the epididymal tissue, dependent on dilutions The
378 number of videos, tracks, and cells analyzed are reported in Table 1.
379 We characterized sperm aggregation by scoring CASA videos using direct observations
380 because the system tracks particles, and thus each track may represent a single cell or an
381 aggregate. We counted the number of cells represented by each track on at least three different
382 frames/track. From these data, we calculated the proportion of cells that aggregated for each male
383 by dividing the total number of aggregates by the total number of motile cells across all tracks
384 (55). Then we calculated the mean number of cells in aggregate (i.e., aggregate size) by dividing
385 the sum of cells in aggregate by the sum of aggregates, both across all tracks for each male and

386 across all males for each species. Finally, we calculated the coefficient of variation (CV) for both
387 the proportion and number of cells that aggregated within each species using the following
388 formula: (standard deviation/mean) x 100. In addition, to further characterize differences in sperm
389 aggregation for the two species whose sperm were observed to aggregate the most extensively, *P.*
390 *maniculatus* and *P. polionotus*, we qualitatively scored the composition and orientation of cells
391 within the sperm aggregates. For males within these species, we calculated the proportion of
392 aggregates that were: 'aligned' in which all sperm adhered to one another in a head-to-flagella
393 orientation and included no immotile, morphologically abnormal or damaged cells, 'defective',
394 which included one or more sperm that was abnormal, immotile, or stuck to the slide, and
395 'opposed' if all sperm are normal and motile, but were not oriented in the same direction, the latter
396 of which included star-shaped aggregates (28).

397 We recorded the following metrics for each recorded track (i.e., single sperm and aggregates):
398 *straight-line velocity* (VSL; calculated using the straight-line distance between the first and last
399 detected positions of the sperm head, divided by the time taken to swim the track; also known as
400 average velocity), *curvilinear velocity* (VCL; calculated using the summed distance between the
401 sperm head positions in each frame divided by the time taken to travel the track; also known as
402 speed), *average path velocity* (VAP; the time-averaged velocity of the sperm head along its
403 average path), and *linearity* (LIN; the ratio of VSL to VCL to measure the straightness of the
404 trajectory; 56). We calculated the mean of each kinematic parameter for both single cells and
405 sperm aggregates separately for each of three populations of sperm cells: all cells, motile cells
406 (i.e., devoid of visually inspected tracks in which cells were unmoving, stuck, or featured an
407 obvious morphological abnormality such as a kinked midpiece), and progressively motile cells
408 (i.e., motile cells with a VSL $\geq$25μm/sec). We used these data to calculate the proportion of motile
409 aggregates by dividing the sum of motile aggregates by the total number of sperm aggregates, and
410 the proportion of progressively motile aggregates by dividing the sum of progressive aggregates
411 by the total number of sperm aggregates for each male (55). For our kinematic analyses, we

412 focused on the motile sperm dataset to eliminate artifacts from damaged or dead cells, and the total
413 sperm population to quantify frequency of aggregation and aggregate size in an effort to capture
414 natural aggregation rates (results from other cell populations are reported in Table S1, Figure S1).
415
416 (c) Statistical analyses
417 We performed all statistical analyses using R version 3.4.2 (57) and visually inspected
418 diagnostic plots (qqplots and plots of the distribution of the residuals against fitted values) to
419 validate model normality. Only the best fitting models are reported here. We created all figures
420 using the 'ggplot2' package with R (58). One *P. californicus* male was excluded from the
421 aggregate analysis dataset because their measurements represented clear outliers. All means are
422 presented ± 1 standard error.
423 To compare species differences in the proportion of aggregated cells, we used the mean values
424 for each male and a generalized linear mixed model (GLMM) using the glmer function from the
425 "lme4" R package and a logit link function (59). The binomial response was the number of sperm
426 cells in aggregate, and the total number of sperm cells was the binomial denominator. In the initial
427 statistical model, we observed the residual deviance to be larger than the residual degrees of
428 freedom, which is an indication of overdispersion (55). We thus used an observation-level random
429 effect (OLRE) as a random factor in all subsequent analyses to control for overdispersion (60). We
430 considered family ID as a random factor in the initial model and both random factors were then
431 used in bivariate analyses for predictors of interest that could potentially explain differences in the
432 proportion of aggregated cells. These predictors included male age, pairing status, the timing of
433 video recordings relative to harvest of the epididymal tissue, and the number of videos recorded.
434 Only predictors that had a *p*-value at or below 0.20 were considered for the final model. We
435 further screened these predictors for collinearity with other significant predictors using linear
436 models and removed collinear predictors, so that only the one with greater relative significance
437 was included in the final GLMM. The remaining model included pairing status and species as

438 fixed factors. We calculated the variance inflation factor (VIF) values and found evidence of
439 collinearity due to two VIF values above the recommended threshold of three (61). Thus, we
440 removed pairing status as a fixed factor and family ID as a random factor, leaving only species as
441 a fixed factor within our final model. Post-hoc pairwise comparisons were performed using Tukey
442 HSD adjustments for multiple comparisons from the "LSmeans" R package (62).

443 To compare species differences in the number of aggregated cells, we used the mean values
444 for each male and initially used a linear mixed model (LMM) using the lmer function from the
445 "lme4" R package, but eventually reverted to using a linear model (LM) because the family ID
446 random factor did not significantly contribute to the residual variability in the response variable.
447 Predictors that were considered for the initial LMM included male age, pairing status, the timing
448 of video recordings relative to harvest of the epididymal tissue, total sperm cells, number of
449 recorded videos, an interaction between the latter two variables, and the ratio of total sperm cells
450 to the number of videos recorded. We considered predictors with p-value<0.20 for the final model,
451 but first screened each for collinearity with other significant predictors using a linear model and
452 removed whenever collinearity was present so that only the one with the greater relative
453 significance was included in the final LM. We dropped non-significant explanatory variables one
454 at a time based on model comparisons using an analyses of variance test to determine the minimal
455 adequate model, but were unable to meet the normality assumptions for this model. We also
456 assessed species differences in the proportion of aggregates that were motile or progressively
457 motile from the total population of aggregates using binomial GLMM. The binomial response was
458 the number of motile sperm aggregates, and the total number of sperm aggregates was the
459 binomial denominator. Our final model contained both an OLRE due to detected overdispersion
460 and family ID as random factors.

461 To determine if sperm aggregates have motility or force benefits over single cells, we initially
462 performed a principal component analysis (PCA) using three related swimming performance
463 measures (VSL, VCL, and VAP) to reduce dimensionality and obtain a composite measure for

464 motile solitary and aggregated sperm cells in both low- and high-viscosity media. Because we
465 found that species distributions overlapped for this composite measure for both single sperm and
466 aggregates in low- and high-viscosity media (Figures S2 and S3), we focused on individual
467 kinematic parameters (VSL, VCL, VAP, LIN) within each species separately. Using our dataset of
468 mean values per male, we used a paired student's t-test to compare each kinematic parameter
469 between solitary sperm cells and sperm aggregates within males for each separate species to
470 determine if sperm aggregates have motility benefit over single cells. To determine if aggregates
471 have a force benefit over single sperm, we conducted these same analyses in high-viscosity media.
472 We then combined these two datasets in low- and high- viscosity media and compared sperm
473 aggregate kinematics in both low- and high- viscosity media at the intra-male level using a paired
474 student's t-test within each species.

475 To assess the structure of *P. maniculatus* and *P. polionotus* aggregates, we used generalized
476 linear models (GLM) to compare the proportions of 'defective and 'opposed' aggregates to
477 'aligned' aggregates. For the composition GLM, the binomial response was the number of
478 'defective' sperm aggregates, and the total number of sperm aggregates that were either 'defective'
479 or 'aligned' was the binomial denominator. We used a paired student's t-test to compare the VCL
480 of these aggregates within males. For the orientation GLM, the binomial response was the number
481 of 'opposed' sperm aggregates, and the total number of sperm aggregates that were either
482 'opposed' or 'aligned' was the binomial denominator. We used a paired student's t-test to compare
483 the VCL of these aggregates within males.

484 Finally, to account for variation in phylogenetic relationships among of the species used in this
485 study, we adopted a phylogenetic generalized least squares approach (63, 64) using the "caper"
486 (65) and "APE" (66) packages in R and using an ultra-metric phylogenetic tree of *Peromyscus*
487 (provided by Dr. Roy Neal Platt II, Texas Biomedical Research Institute), based on sequence
488 variation in the mitochondrial gene, cytochrome B. The species' relationships within this tree
489 matched those from other previously established phylogenies of *Peromyscus* (34, 67). We used

490 this phylogeny as a covariate in regression analyses to investigate the effect of relative testis
491 weight on sperm aggregation, including the proportion of aggregated cells and aggregate size, and
492 the within-species CV for each of these parameters. Finally, to control for differences in male
493 body when examining testis mass, we included body mass as a separate fixed factor within our
494 analyses, a method better suited to estimating relative testis weight is size rather than using the
495 ratio of testis to body mass or residuals (68, 69).
496

497 **Acknowledgments:**
498 We are grateful to Erica Glasper for providing *P. californicus* males and Hopi Hoekstra for
499 providing *P. gossypinus* males. Thanks to Harrison Arsis, Madeline-Sophie Dang, Catherine Liu,
500 and Audrey Mvemba for their assistance with video analysis and to Philip Johnson, Danielle
501 Adams, Sam Church, and Shelby Wilson for statistical advice. We thank Neal Platt for providing
502 the ultra-metric *Peromyscus* tree for use in our statistical analysis.
503

504 **References:**

505 1. S. Pitnick, D. J. Hosken, T. R. Birkhead, "Sperm Morphological Diversity" in Sperm Biology,
506 An Evolutionary Perspective, T. R. Birkhead, D. J. Hosken, S. Pitnick, Eds. (Academic Press,
507 2009), pp. 69–149.
508 2. D. M. Higginson, S. Pitnick, Evolution of intra-ejaculate sperm interactions: do sperm
509 cooperate? *Biological Reviews* 86, 249–270 (2011).
510 3. S. Immler, Sperm competition and sperm cooperation: the potential role of diploid and haploid
511 expression. *Reproduction* 135, 275–283 (2008).
512 4. H. D. Moore, K. Dvoráková, N. Jenkins, W. Breed, Exceptional sperm cooperation in the
513 wood mouse. *Nature* 418, 174–177 (2002).
514 5. S. Immler, H. D. Moore, W. Breed, T. Birkhead, By hook or by crook? Morphometry,
515 competition and cooperation in rodent sperm. *PloS one* 2, e170 (2007).

516 6. T. R. Birkhead, A. P. Møller, W. J. Sutherland, Why do females make it so difficult for males
517 to fertilize their eggs? *Journal of Theoretical Biology* 161, 51–60 (1993).
518 7. T. Pizzari, K. R. Foster, Sperm sociality: cooperation, altruism, and spite. *PLoS Biology* 6
519 (2008).
520 8. T. R. Birkhead, J. G. Martinez, T. Burke, D. P. Froman, Sperm mobility determines the
521 outcome of sperm competition in the domestic fowl. *Proceedings of the Royal Society of*
522 *London. Series B: Biological Sciences* 266, 1759–1764 (1999).
523 9. M. J. G. Gage, C. P. Macfarlane, S. Yeates, R. G. Ward, J. B. Searle, G. A. Parker,
524 Spermatozoal traits and sperm competition in Atlantic salmon: relative sperm velocity is the
525 primary determinant of fertilization success. *Current Biology* 14, 44–47 (2004).
526 10. H. S. Fisher, H. E. Hoekstra, Competition drives cooperation among closely related sperm of
527 deer mice. *Nature* 463, 801 (2010).
528 11. M. Pearcy, N. Delescaille, P. Lybaert, S. Aron, Team swimming in ant spermatozoa. *Biology*
529 *Letters* 10, 20140308 (2014).
530 12. F. Hayashi, Sperm co-operation in the fishfly, *Parachauliodes japonicus*. *Functional Ecology*
531 12, 347–350 (1998).
532 13. D. M. Woolley, R. F. Crockett, W. D. I. Groom, S. G. Revell, A study of synchronisation
533 between the flagella of bull spermatozoa, with related observations. *Journal of Experimental*
534 *Biology* 212, 2215–2223 (2009).
535 14. S. Ishijima, S. A. Ishijima, B. A. Afzelius, Movement of Turritella spermatozoa: Direction of
536 propagation and chirality of flagellar bends. *Cell Motility* 44, 85–95 (1999).
537 15. C. K. Tung, C. Lin, B. Harvey, A.G. Fiore, F. Ardon, M. Wu, S. S. Suarez, Fluid
538 viscoelasticity promotes collective swimming of sperm. *Scientific Reports* 7, 1–9 (2017).
539 16. H. D. M. Moore, D. A. Taggart, Sperm pairing in the opossum increases the efficiency of
540 sperm movement in a viscous environment. *Biol Reprod* 52, 947–953 (1995).

541 17. S. S. Suarez, Mammalian sperm interactions with the female reproductive tract. *Cell and*
542 *Tissue Research* 363, 185–194 (2016).
543 18. F. W. Hanson, J. W. Overstreet, The interaction of human spermatozoa with cervical mucus in
544 vivo. *American Journal of Obstetrics and Gynecology* 140, 173–178 (1981).
545 19. D. M. Phillips, Ultrastructure of spermatozoa of the wooly opossum *Caluromys philander*.
546 *Journal of Ultrastructure Research* 33, 381–397 (1970).
547 20. J. M. Bedford, J. C. Rodger, W. G. Breed, Why so many mammalian spermatozoa - a clue
548 from marsupials? *Proceedings of the Royal Society of London B: Biological Sciences* 221,
549 221–233 (1984).
550 21. D. M. Higginson, K. R. H. Henn, Effects of sperm conjugation and dissociation on sperm
551 viability in vitro. *PloS one* 7, e34190 (2012).
552 22. K. S. K. Tung, A. Okada, R. Anagimachi, Sperm autoantigens and fertilization. I. effects of
553 antisperm autoantibodies on rouleaux formation, viability, and acrosome reaction of guinea
554 pig spermatozoa. *Biology of Reproduction* 23, 877–886 (1980).
555 23. J. B. Mackie, M. H. Walker, A study of the conjugate sperm of the dytiscid water beetles
556 *Dytiscus marginalis* and *Colymbetes fuscus*. *Cell Tissue Res.* 148, 505–519 (1974).
557 24. S. M. McGlinn, B. A. Shepherd, J. Martan, Rouleau Formation and fertility of spermatozoa in
558 guinea pigs treated with cyproterone acetate. *Archives of Andrology* 2, 197–202 (1979).
559 25. D. A. Taggart, J. L. Johnson, H. P. O'Brien, H. D. M. Moore, Why do spermatozoa of the
560 American marsupial form pairs? A clue from the analysis of sperm-pairing in the epididymis
561 of the grey short-tailed opossum, *Monodelphis domestica*. *Anatomical Record* 236, 465–478
562 (1993).
563 26. B. Nixon, H. Ecroyd, J. L. Dacheux, F. Dacheux, V. Labas, S. D. Johnston, R. C. Jones,
564 Formation and dissociation of sperm bundles in monotremes. *Biology of Reproduction* 95
565 (2016).

27. D. J. G. Pearce, L. A. Hoogerbrugge, K. A. Hook, H. S. Fisher, L. Giomi, Cellular geometry controls the efficiency of motile sperm aggregates. *Journal of The Royal Society Interface* 15, 20180702 (2018).

28. H. S. Fisher, L. Giomi, H. E. Hoekstra, L. Mahadevan, The dynamics of sperm cooperation in a competitive environment. *Proceedings of the Royal Society of London B: Biological Sciences* 281, 20140296 (2014).

29. D. A. Dewsbury, Interactions between males and their sperm during multi-male copulatory episodes of deer mice (*Peromyscus maniculatus*). *Animal Behaviour* 33, 1266–1274 (1985).

30. D. A. Birdsall, D. Nash, Occurrence of successful multiple insemination of females in natural populations of deer mice (*Peromyscus maniculatus*). *Evolution* 27, 106–110 (1973).

31. D. A. Dewsbury, An exercise in the prediction of monogamy in the field from laboratory data on 42 species of muroid rodents. *Biologist* 63, 138–162 (1981).

32. D. W. Foltz, Genetic evidence for long-term monogamy in a small rodent, *Peromyscus polionotus*. *The American Naturalist* 117, 665–675 (1981).

33. N. L. Bedford, H. E. Hoekstra, The natural history of model organisms: Peromyscus mice as a model for studying natural variation. *eLife Sciences* 4, e06813 (2015).

34. L. M. Turner, *et al.*, Monogamy evolves through multiple mechanisms: evidence from V1aR in deer mice. *Molecular Biology and Evolution* 27, 1269–1278 (2010).

35. S. A. Ramm, G. A. Parker, P. Stockley, Sperm competition and the evolution of male reproductive anatomy in rodents. *Proceedings of the Royal Society B: Biological Sciences* 272, 949–955 (2005).

36. M. Tourmente, D. Zarka-Trigo, E. R. S. Roldan, Is the hook of muroid rodent's sperm related to sperm train formation? *Journal of Evolutionary Biology* 29, 1168–1177 (2016).

37. D. M. Higginson, K. B. Miller, K. A. Segraves, S. Pitnick, Convergence, recurrence and diversification of complex sperm traits in diving beetles (Dytiscidae). *Evolution* 66, 1650–1661 (2012).

592 38. G. A. Parker, Why are there so many tiny sperm? Sperm competition and the maintenance of
593 two sexes. *Journal of Theoretical Biology* 96, 281–294 (1982).
594 39. R. R. Snook, Sperm in competition: not playing by the numbers. *Trends in Ecology &*
595 *Evolution* 20, 46–53 (2005).
596 40. S. Calhim, S. Immler, T. R. Birkhead, Postcopulatory sexual selection is associated with
597 reduced variation in sperm morphology. *PLoS ONE* 2, e413 (2007).
598 41. J. L. Fitzpatrick, B. Baer, Polyandry reduces sperm length variation in social insects. *Evolution*
599 65, 3006–3012 (2011).
600 42. S. Immler, S. Calhim, T. R. Birkhead, Increased postcopulatory sexual selection reduces the
601 intramale variation in sperm design. *Evolution* 62, 1538–1543 (2008).
602 43. L. Carballo, A. Battistotti, K. Teltscher, M. Lierz, A. Bublat, M. Valcu, B. Kempenaers,
603 Sperm morphology and evidence for sperm competition among parrots. *Journal of*
604 *Evolutionary Biology* 32, 856–867 (2019).
605 44. A. Rowley, L. Locatello, A. Kahrl, M. Rego, A. Boussard, E. Garza-Gisholt, R. M. Kempster,
606 S. P. Collin, E. Giacomello, M. C. Follesa, C. Porcu, Sexual selection and the evolution of
607 sperm morphology in sharks. *Journal of Evolutionary Biology* 32, 1027–1035 (2019).
608 45. Y. Takami, T. Sota, Sperm competition promotes diversity of sperm bundles in *Ohomopterus*
609 ground beetles. *Naturwissenschaften* 94, 543–550 (2007).
610 46. D. M. Higginson, K. B. Miller, K. A. Segraves, S. Pitnick, Female reproductive tract form
611 drives the evolution of complex sperm morphology. *Proceedings of the National Academy of*
612 *Sciences* 109, 4538–4543 (2012).
613 47. W. Eberhard, *Female Control: Sexual Selection by Cryptic Female Choice* (Princeton
614 University Press, 1996).
615 48. I. F. Greenbaum, R. J. Baker, P. R. Ramsey, Chromosomal evolution and the mode of
616 speciation in three species of Peromyscus. *Evolution* 32, 646–654 (1978).

49. R. C. Firman, L. W. Simmons, Experimental evolution of sperm competitiveness in a mammal. *BMC Evolutionary Biology* 11 (2011).
50. J. E. Simons, S. D. Olson, "Sperm motility: models for dynamic behavior in complex environments" in Cell Movement, M. Stolarska, N. Tarfulea, Eds. (Springer International Publishing, 2018), pp. 169–209.
51. Y. Ishikawa, T. Usui, M. Yamashita, Y. Kanemori, T. Baba, Surfing and swimming of ejaculated sperm in the mouse oviduct. *Biology of Reproduction* 94, 89, 1–9 (2016).
52. S. Wang, I. V. Larina, In vivo three-dimensional tracking of sperm behaviors in the mouse oviduct. *Development* 145, dev157685 (2018).
53. V. Magdanz, S. Boryshpolets, C. Ridzewski, B. Eckel, K. Reinhardt, The swim-up technique separates bovine sperm by metabolic rates, motility and tail length. *bioRxiv*, 624502 (2019).
54. S. S. Suarez, X. Dai, Hyperactivation enhances mouse sperm capacity for penetrating viscoelastic media. *Biology of Reproduction* 46, 686–691 (1992).
55. M. J. Crawley, *The R Book Second Edition* (John Wiley & Sons, Ltd., 2013).
56. World Health Organization, "WHO laboratory manual for the examination and processing of human semen" (2010).
57. R Core Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2016).
58. H. Wickham, *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2016).
59. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67, 1–48 (2015).
60. X. A. Harrison, Using observation-level random effects to model overdispersion in count data in ecology and evolution. *PeerJ* 2, e616 (2014).
61. A. F. Zuur, E. N. Ieno, C. S. Elphick, A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution* 1, 3–14 (2010).

642 62. R. V. Lenth, Least-squares means: The R package lsmeans. *Journal of Statistical Software* 69,
643 1–33 (2016).
644 63. M. Pagel, Inferring the historical patterns of biological evolution. *Nature* 401, 877–884
645 (1999).
646 64. R. P. Freckleton, P. H. Harvey, M. Pagel, Phylogenetic analysis and comparative data: a test
647 and review of evidence. *The American Naturalist* 160, 712–726 (2002).
648 65. D. Orme, R. Freckleton, G. Thomas, T. Petzoldt, The caper package: comparative analysis of
649 phylogenetics and evolution in R. *R package version* 5, 1–36 (2013).
650 66. E. Paradis, J. Claude, K. Strimmer, APE: Analyses of phylogenetics and evolution in R
651 language. *Bioinformatics* 20, 289–290 (2004).
652 67. R. D. Bradley, N.D. Durish, D. S. Rogers, J. R. Miller, M. D. Engstrom, C. W. Kilpatrick,
653 Toward a molecular phylogeny for Peromyscus: evidence from mitochondrial cytochrome- *b*
654 sequences. *Journal of Mammalogy* 88, 1146–1159 (2007).
655 68. E. García-Berthou, On the misuse of residuals in ecology: testing regression residuals vs. the
656 analysis of covariance. *Journal of Animal Ecology* 70, 708–711 (2001).
657 69. S. Lüpold, G. M. Linz, J. W. Rivers, D. F. Westneat, T. R. Birkhead, Sperm competition
658 selects beyond relative testes size in birds. *Evolution* 63, 391–402 (2009)

659
660
661
662
663
664
665

FIGURES AND TABLES

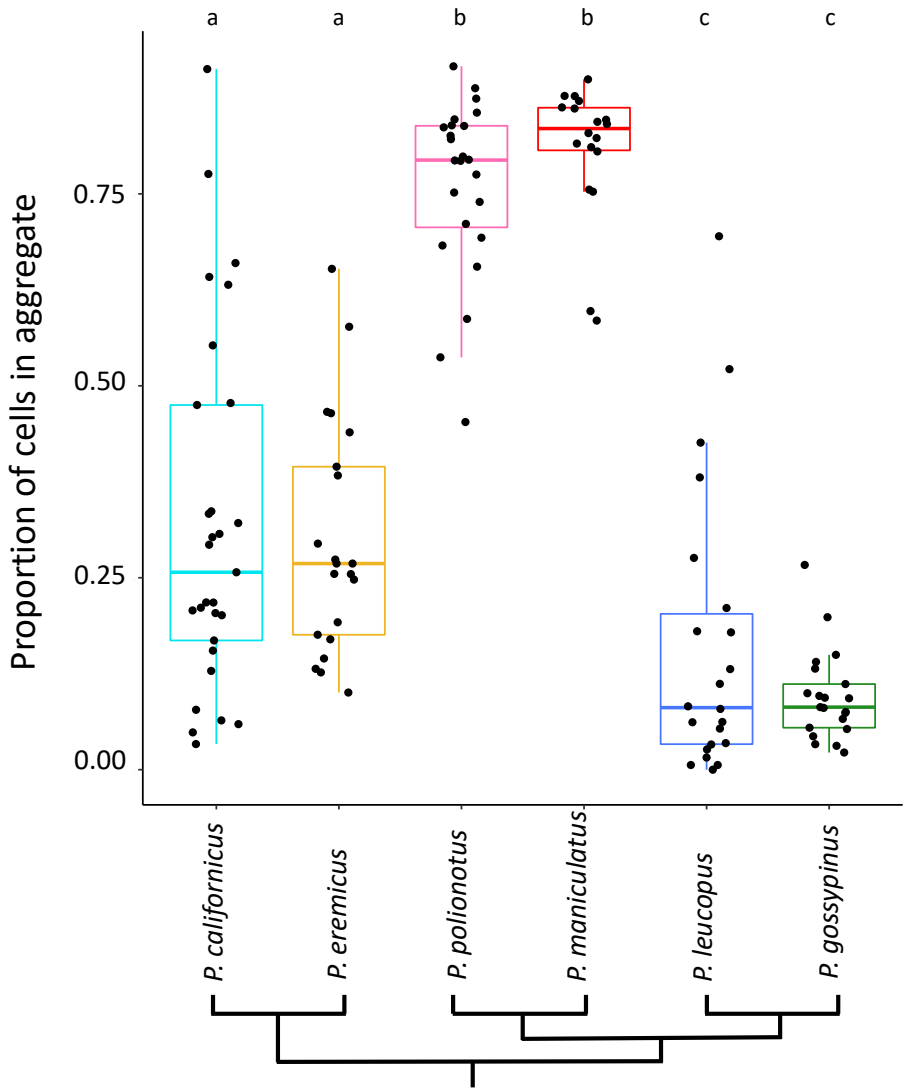

Figure 1.

The observed natural variation in the proportion of aggregated sperm cells for six closely related species of *Peromyscus* mice (phylogeny adapted from Bradley et al. 2007). Box-plots represent median and interquartile ranges with raw data overlaid. Statistically significant differences at the $P = 0.05$ level are denoted by differing letters; shared letters denote no statistical difference.

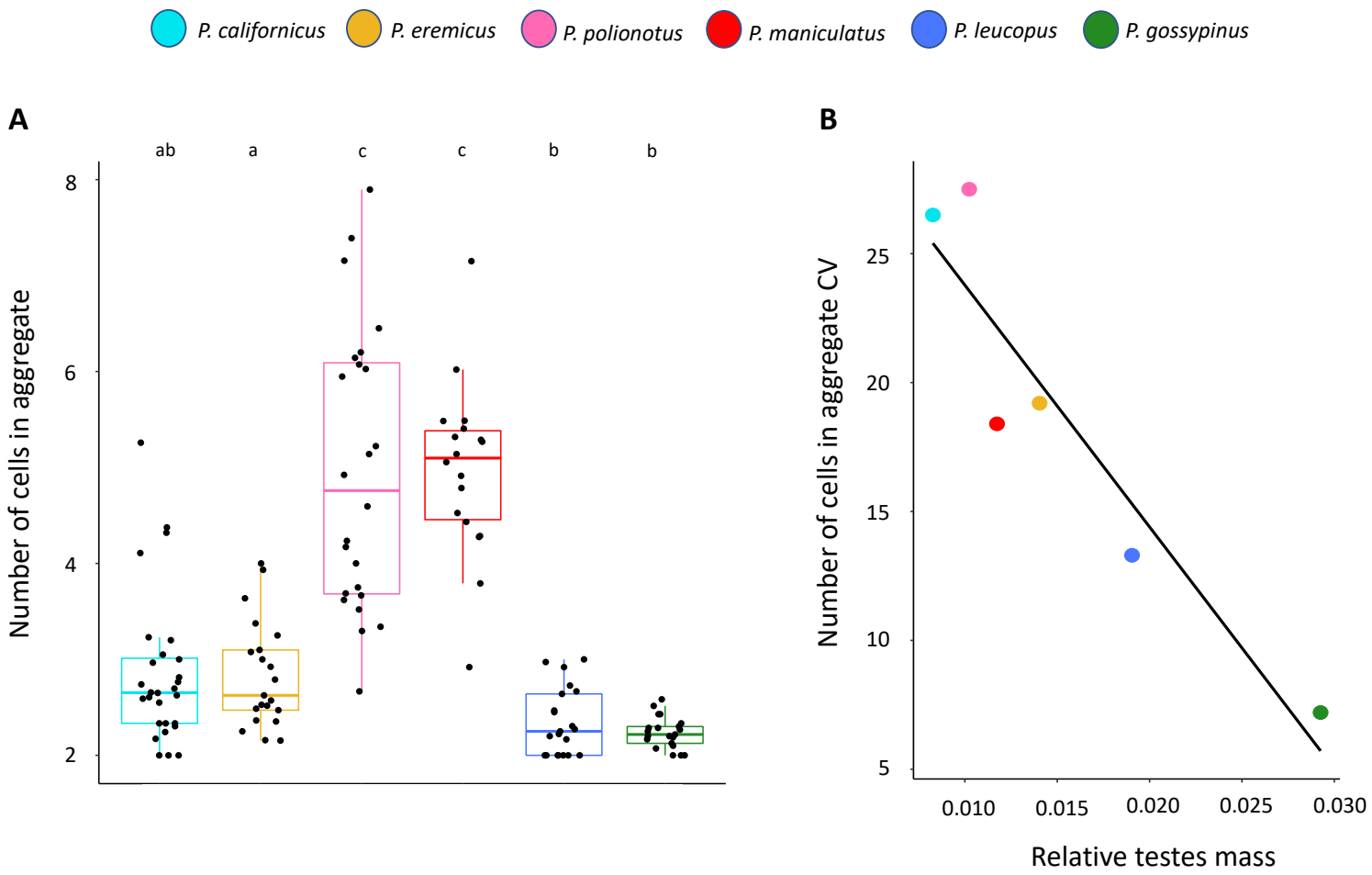

Figure 2.

Sperm aggregate size varies among species and is negatively associated with relative testes size. (A) The observed natural variation in the number of aggregated sperm cells for six closely related species of *Peromyscus* mice. Box-plots represent median and interquartile ranges with raw data overlaid. Statistically significant differences at the $P = 0.05$ level are denoted by differing letters; shared letters denote no statistical difference. (B) When controlling for phylogenetic relationships, the coefficient of variation (CV) for the number of aggregated sperm cells negatively correlates with relative testis mass across these species. Note truncated y-axes.

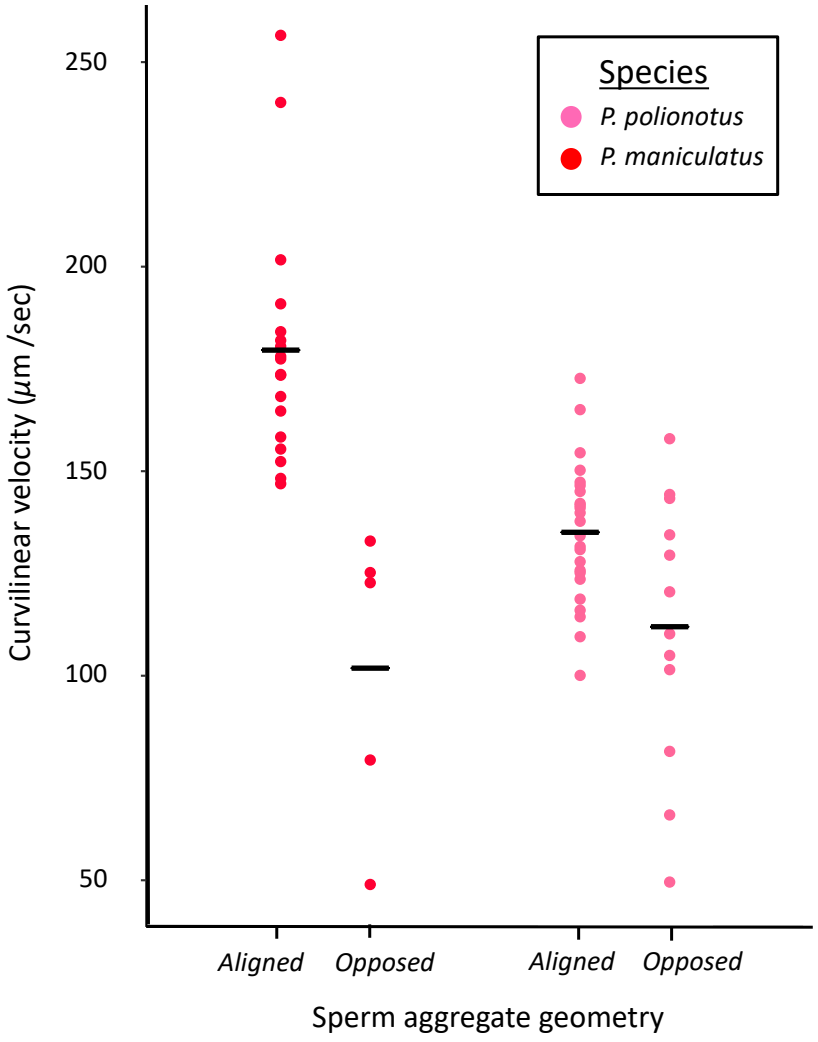

Figure 3.

The effect of sperm aggregate geometry on the curvilinear velocity (μm/sec) of sperm aggregates for two species that aggregated most – *Peromyscus maniculatus* and *Peromyscus polionotus*. Circles represent mean values per male within each species, and black lines represent the mean value within each category. Note truncated y-axis.

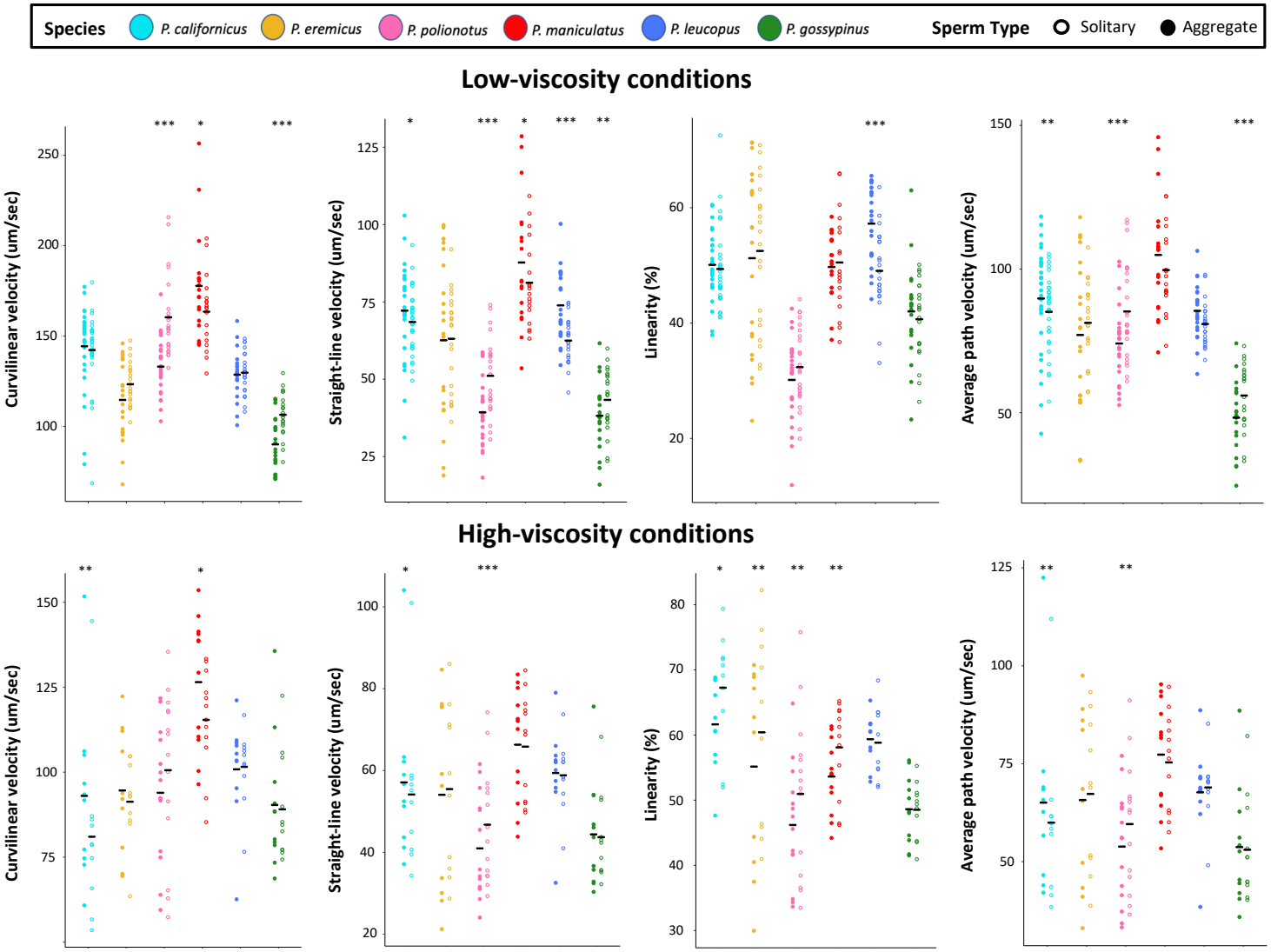

Figure 4.

Kinematic parameters of sperm aggregates (closed circles) and solitary sperm cells (open circles) for six species of *Peromyscus* mice in low- and high-viscosity conditions. Circles represent mean values per male, and black lines represent species means. Statistical significance levels comparing aggregated and solitary cells within each species are indicated by * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. Note truncated y-axes.

TABLE 1.

Summary of live sperm aggregate results for mice in the genus *Peromyscus*

Peromyscus Species	No. males	No. sperm	% motile sperm	% progressive sperm	No. sperm aggregates	% cells in aggregate	% motile aggregates	% progressive aggregates
californicus	28	8235	81.2%	78.9%	823	29.7% (2446/8235)	90.5% (745/823)	86.6% (713/823)
eremicus	21	4906	67.4%	58.2%	513	30.2% (1481/4906)	77.8% (399/513)	66.7% (342/513)
polionotus	24	6360	82.4%	57.3%	949	79.5% (5059/6360)	90.9% (863/949)	66.1% (627/949)
maniculatus	18	4991	85.0%	84.3%	822	81.3% (4059/4991)	93.4% (768/822)	92.3% (759/822)
leucopus	22	6341	87.6%	84.5%	361	15.0% (949/6341)	90.3% (326/361)	88.1% (318/361)
gossypinus	21	5970	82.2%	70.3%	254	9.6% (576/5970)	88.6% (225/254)	70.1% (178/254)

Motile sperm cells include only those that exhibited movement patterns. Progressive sperm cells are motile cells with a straight-line velocity $\geq$ 25 μm/sec.

TABLE 2.

Fixed effects from a binomial generalized linear mixed model examining differences in the proportion of aggregated sperm cells and a linear model examining differences in the number of aggregated sperm cells across six species of *Peromyscus* mice

GLMM: PROPORTION OF CELLS IN AGGREGATE					
Model Term	Beta (SE)	Exp (beta)	95% CI	z	Pr(>|z|)
Intercept	1.52 (0.22)				
eremicus	-2.47 (0.30)	0.07	(0.04, 0.13)	-8.14	**< 0.001**
gossypinus	-3.94 (0.31)	0.02	(0.01, 0.03)	-12.89	**< 0.001**
californicus	-2.58 (0.29)	0.07	(0.04, 0.12)	-9.06	**< 0.001**
leucopus	-3.79 (0.30)	0.02	(0.01, 0.04)	-12.46	**< 0.001**
polionotus	-0.25 (0.29)	0.44	(0.30, 0.58)	-0.85	0.393
LM: NUMBER OF CELLS IN AGGREGATE					
Model Term	Beta (SE)	Exp (beta)	95% CI	t	Pr(>|z|)
Intercept	3.90 (0.26)				
Total Sperm Cells	0.00 (0.00)	0.50	(0.50, 0.50)	5.83	**< 0.001**
eremicus	-1.97 (0.24)	0.12	(0.08, 0.18)	-8.15	**< 0.001**
gossypinus	-2.77 (0.24)	0.06	(0.04, 0.09)	-11.53	**< 0.001**
californicus	-2.19 (0.23)	0.10	(0.07, 0.15)	-9.68	**< 0.001**
leucopus	-2.68 (0.24)	0.06	(0.04, 0.10)	-11.18	**< 0.001**
polionotus	0.04 (0.23)	0.51	(0.40, 0.62)	0.16	0.877

For both models, all rows are being compared with the intercept – *Peromyscus maniculatus*. 95% confidence intervals (CI) were calculated for each effect size.

TABLE 3.

Results from an intra-male analysis comparing motile solitary and aggregated sperm kinematics in low- and high-viscosity conditions for six species of *Peromyscus* mice to test whether sperm aggregates confer kinematic benefits (shaded in gray)

PEROMYSCUS SPECIES	df	KINEMATIC VARIABLE			
		Curvilinear Velocity (μm/sec)	Linearity (VSL/VCL)	Straight-Line Velocity (μm/sec)	Average Path Velocity (μm/sec)
LOW-VISCOSITY CONDITIONS					
californicus	*28*	t = -1.0545, p = 0.3007	t = -0.56153, p = 0.5789	t = -2.2982, **p = 0.02923**	t = - 2.8162, **p = 0.008805**
eremicus	*20*	t = 1.6225, p = 0.1204	t = 0.73517, p = 0.4708	t = 0.1312, p = 0.8969	t = 1.0464, p = 0.3079
polionotus	*23*	t = 9.4575, **p = 2.1566e-09**	t = 1.699, p = 0.1028	t = 5.8355, **p = 6.026e-06**	t = 6.0729, **p = 3.408e-06**
maniculatus	*17*	t = 2.2482, **p = 0.03812**	t = 0.48075, p = 0.6368	t = -2.2335, **p = 0.03924**	t = -1.9206, p = 0.07172
leucopus	*20*	t = 0.23337, p = 0.8178	t = -4.8385, **p = 9.973e-05**	t = -5.5521, **p = 1.959e-05**	t = -1.838, p = 0.08096
gossypinus	*20*	t = 5.4048, **p = 2.73e-05**	t = -1.0247, p = 0.3177	t = 3.715, **p = 0.001369**	t = 4.479, **p = 0.0002298**
HIGH-VISCOSITY CONDITIONS					
californicus	*9*	t = -3.7465, **p = 0.003357**	t = 2.682, **p = 0.02512**	t = -2.7743, **p = 0.0216**	t = -4.1106, **p = 0.002634**
eremicus	*9*	t = -0.71739, p = 0.4913	t = 4.2087, **p = 0.002277**	t = 0.62918, p = 0.5449	t = 0.56505, p = 0.5858
polionotus	*13*	t = 2.0349, p = 0.06278	t = 3.9242, **p = 0.001745**	t = 4.8314, **p = 0.0003279**	t = 4.18, **p = 0.001079**
maniculatus	*11*	t = -2.9397, **p = 0.01345**	t = 3.9259, **p = 0.002369**	t = -0.30835, p = 0.7636	t = -1.0543, p = 0.3114
leucopus	*9*	t = 0.32429, p = 0.7531	t = -0.31635, p = 0.759	t = -0.34159, p = 0.7405	t = 0.95055, p = 0.3667
gossypinus	*10*	t = -0.63835, p = 0.95376	t = -0.076451, p = 0.9406	t = -0.60402, p = 0.5593	t = -0.71643, p = 0.4901

Statistical results are based on paired-student t-tests.