

1 **Rapid evolution of coordinated and collective movement in response to artificial selection**

2

3 Short title: **Evolution of social interaction rules**

4

5 Alexander Kotrschal<sup>\*1,2†</sup>, Alexander Szorkovszky<sup>1,3†</sup>, James Herbert-Read<sup>4,5</sup>, Natasha I. Bloch<sup>6</sup>,

6 Maksym Romenskyy<sup>7</sup>, Séverine Denise Buechel<sup>1</sup>, Ada Fontrodona Eslava<sup>1,8</sup>, Laura Sánchez Alòs<sup>1</sup>,

7 Hongli Zeng<sup>9</sup>, Audrey Le Foll<sup>1</sup>, Ganaël Braux<sup>1</sup>, Kristiaan Pelckmans<sup>10</sup>, Judith E. Mank<sup>11,12</sup>, David

8 Sumpter<sup>3</sup> & Niclas Kolm<sup>1</sup>

9 † equal contribution

10 \* corresponding author: [alexander.kotrschal@zoologi.su.se](mailto:alexander.kotrschal@zoologi.su.se)

11

12 **Affiliations**

13 <sup>1</sup> Stockholm University, Department of Zoology / Ethology, Stockholm, Sweden

14 <sup>2</sup> Wageningen University, Behavioural Ecology, Wageningen, Netherlands

15 <sup>3</sup> Uppsala University, Department of Mathematics, Uppsala, Sweden

16 <sup>4</sup> Department of Zoology, University of Cambridge, Cambridge, UK

17 <sup>5</sup> Aquatic Ecology, Lund University, Lund, Sweden

18 <sup>6</sup> University of Los Andes, Department of Biomedical Engineering, Bogotá, Colombia

19 <sup>7</sup> Imperial College London, Department of Life Sciences, London, United Kingdom

20 <sup>8</sup> University of St. Andrews, Centre for Biological Diversity, St. Andrews, United Kingdom

21 <sup>9</sup> Nanjing University of Posts and Telecommunications, School of Science, Nanjing, China

22 <sup>10</sup> Uppsala University, Department of Information Technology, Uppsala, Sweden

23 <sup>11</sup> University College London, London, United Kingdom

24 <sup>12</sup> University of British Columbia, Department of Zoology, Vancouver, Canada

25

26

## 27 **Abstract**

28 Collective motion occurs when individuals use social interaction rules to respond to the movements  
29 and positions of their neighbors. How readily these social decisions are shaped by selection remains  
30 unknown. Through artificial selection on fish (guppies, *Poecilia reticulata*) for increased social  
31 coordination (group polarization), we demonstrate that social interaction rules can evolve  
32 remarkably fast. Within just three generations, groups of polarization selected females showed a  
33 15% increase in polarization, coupled with increased cohesiveness, compared to fish from control  
34 lines. They did not differ in physical swimming ability or exploratory behavior. However,  
35 polarization selected fish adopted faster speeds, particularly in social contexts, and showed stronger  
36 alignment and attraction responses to multiple neighbors. Our results demonstrate that animals'  
37 social interactions can rapidly evolve under strong selection, and reveal which social interaction  
38 rules change when collective behavior evolves.

## 39 **Introduction**

40 Moving animal groups display spectacular forms of coordinated behavior, with individuals moving  
41 together with high degrees of spatial and directional organization. This organization is often  
42 achieved by individuals using interaction 'rules' to respond to their neighbors' movements and  
43 positions. For example, attraction, repulsion and alignment responses can act to maintain the  
44 cohesiveness and directional organization of groups (1, 2). The details of these interactions, and the  
45 social information individuals use to inform these decisions are now well described across many  
46 species (3-7). However, despite our growing knowledge of the mechanistic nature of social  
47 interactions in moving animal groups, we still know very little about the evolution of these social  
48 rules (8, 9).

49 For instance, while it has been established that intraspecific variation exists in animals' social  
50 attraction and alignment towards conspecifics (6, 10, 11), it remains unclear whether such variation  
51 can be attributed to heritable differences in individuals' social behavior, or is instead being driven  
52 by differences in individuals' state, age, experience, or size (12). Indeed, while there are inherited  
53 differences in the tendencies of marine or benthic sticklebacks' to school (13), those differences  
54 appear to be driven by genes affecting how social information is detected by neighbors (genes  
55 affecting the lateral lines system), and not necessarily how that information is behaviorally acted  
56 upon. Nevertheless, evolutionary models suggest that heritable differences in social decision-  
57 making should exist and persist in populations (14), and particular environments should favor  
58 particular social interactions depending on the selective forces present (15). What kinds of  
59 interactions are subject to selection, however, remains unclear.

60 In order to determine how selection can shape the social interaction rules that animals use to  
61 coordinate their movements, we performed a four-year artificial selection experiment using the  
62 guppy (*Poecilia reticulata*). Guppies are a model species for social behavior and evolution (16), and  
63 although they naturally shoal (17), their schooling tendencies tend to be weaker than in other  
64 species of fish, offering the potential for selection to increase social coordination. Our selection  
65 procedure targeted group polarization, a standard measure of directional coordination in animal  
66 groups. This metric captures the tendency of group members to align with each other's directional  
67 headings. By artificially selecting for polarization over multiple generations we tested whether, and  
68 how quickly, coordinated group movement evolved when under strong directional selection.  
69 Importantly, polarization can only be measured in a group context but we nevertheless could apply  
70 an individual-level selection approach; our recently developed sorting protocol of repeated mixing  
71 and polarization-determination concentrates the individuals with the highest polarization  
72 propensities in few groups (18, 19). Those individuals could then be bred for the selection lines.  
73 Our artificial selection approach further allowed us to measure how selection shaped the social rules  
74 responsible for increased polarization in these groups.

75 Based on previous simulation and empirical studies, we had a number of *a priori* candidate  
76 mechanisms for how increased polarization could be achieved. These mechanisms include increased  
77 strength of alignment or attraction responses (20, 21), increased interaction ranges (22), increased  
78 number of influential neighbors (23), more frequent directional updating (24), faster speeds in  
79 social contexts (5, 25), or changes to individuals' exploration or boldness (10). Here we identify  
80 which of these changes occurred to individuals' social interaction rules following selection.

## 81 **Results**

### 82 **The Artificial Selection Procedure**

83 Across three independent selection lines (i.e.  $n = 3$  replicate lines), we used a previously validated  
84 sorting method (18, 19) to identify the top 20% of female fish that consistently formed more  
85 polarized groups, and subsequently bred from those individuals. We focused on female behavior in  
86 our selection experiment because females of this species have a higher propensity to shoal than  
87 males (17). The sorting method involved open-field assays on groups of eight female fish ( $n = 16$   
88 groups per replicate), where fish were filmed when they explored an empty circular arena (diameter  
89 550 mm, water depth 3 cm) together for ten minutes. Fish were then tracked using IDTracker (26)  
90 from the second to tenth minute, inclusive, from which the fish' trajectories were subsequently  
91 analyzed. Across all frames in an assay, we calculated a group's polarization (given by the total  
92 length of the sum of the eight unit vectors characterizing the orientation of each fish, divided by  
93 eight). Polarization scores closer to one indicate fish are oriented in the same direction, while scores

94 closer to zero indicate fish are less aligned. After being assayed, the 16 groups were ranked for their  
95 median polarization scores, and half of each group's members were subsequently swapped between  
96 adjacently ranked groups. This ranking and mixing of groups was repeated for 12 rounds, allowing  
97 us to create repeatable variation in polarization between groups (19). Twenty-six females from the  
98 four top-ranked groups in each line were then paired with unsorted males to breed the next  
99 generation of polarization-selected fish. To establish control lines ( $n = 3$ ), we took 26 randomly  
100 selected females from the remaining groups, and bred from those fish. Once the progeny from each  
101 line were fully mature, the sorting method was performed again on the next two generations of  
102 females, providing a total of three generations of selection. Polarization and control line females  
103 were always paired with males from their own cohort. To ensure the control lines experienced the  
104 same experimental conditions as the polarization lines, control fish were placed in arenas and mixed  
105 between groups in the same way as the polarization lines, but were not sorted. For further details of  
106 the selection procedure see Fig. 1 and (18, 19).

### 107 **Evidence for Selection**

108 We performed shoaling assays (as above) on the offspring of the polarization and control lines from  
109 generation three. We found that the polarization of groups across the three replicates was on  
110 average 15% higher in polarization lines ( $n = 88$  groups) compared to control lines ( $n = 85$  groups;  
111 difference replicate 1: 8.4%, replicate 2: 19.7%, replicate 3: 18.7%; LMM for all replicates:  $t =$   
112 6.45,  $df = 170$ ,  $P < 0.001$ ; Fig. 2A). Males did not display a significant response to selection  
113 ( $t=1.13$ ,  $df=109$ ,  $P=0.26$ ), but weak differences between polarization and control lines existed in  
114 other behavioral measures consistent with the females. See Fig. S1 for results over all generations  
115 and discussion of the males.

### 116 **Changes to Individuals' Movement and Behaviour**

117 We next tested whether selection had changed the movement characteristics of the fish in the  
118 polarization compared to control lines. As in many other fish species, guppies move with  
119 intermittent burst and glide phases (15), allowing us to characterize their movements in discrete  
120 steps (Fig. 3A, 3B). Groups of females from the polarization lines exhibited a  $13.5 \text{ mm s}^{-1}$  (26%)  
121 higher median speed in comparison to control lines (LMM:  $t = 5.59$ ,  $df = 170$ ,  $P < 0.001$ ). We also  
122 performed open-arena assays on single fish and found that the difference in median speed between  
123 polarization and control lines was still significant, but less pronounced compared to the social  
124 context. Single fish from polarization lines were on average,  $8.2 \text{ mm s}^{-1}$  (17%) faster than single  
125 fish from control lines (LMM:  $t = 2.04$ ,  $df = 117$ ,  $P = 0.043$ ). As speed is highly correlated with  
126 group polarization in shoaling fish (10), the increase in polarization seen in the selection lines could  
127 have been due to non-social selection for faster moving fish, or reduced swimming abilities in fish

128 from the control lines. However, the polarization-selected lines were still 5.7% more polarized  
129 when controlling for median speed differences between the lines (LMM:  $t = 2.52$ ,  $df = 169$ ,  $P =$   
130  $0.013$ ); and there were no differences between the swimming abilities of fish in the polarization and  
131 control lines when tested for maximal swimming speed and endurance in a swim tunnel (LMM:  $t =$   
132  $-0.56$ ,  $df = 64$ ,  $P = 0.579$ ; Fig. S2). Differences in behavior might also reflect differences between  
133 the polarization and control lines in overall ‘boldness’ or tendency to explore the arena, however,  
134 there were no differences in emergence time (i.e. ‘boldness’; LMM:  $t = -0.12$ ,  $df = 28$ ,  $P = 0.909$ ;  
135 Fig. S3) or exploration (LMM:  $t = -0.38$ ,  $df = 28$ ,  $P = 0.704$ ; Fig. S3) between the polarization and  
136 control lines when tested using a standard assay.

137 In order to further investigate whether the social environment affects the speed that fish adopted in  
138 the polarization and control lines, we identified the speeds at which fish decided to accelerate ( $|v|_{\min}$ ;  
139 Fig. 3A) and plotted this as a function of the distance to their nearest neighbor (Fig. 3C). We found  
140 that while fish from the polarization lines generally maintained higher speeds than fish from control  
141 lines, these differences were particularly apparent when fish were close to their neighbors, with  
142 differences in speed between the lines becoming less pronounced as neighbors moved further apart.  
143 This provides further support that differences in speed were, at least in part, modulated by  
144 interactions with conspecifics. Polarization and speed results were also robust when controlling for  
145 potential differences in thigmotaxis (‘wall-hugging’, i.e. propensity of swimming close to the walls)  
146 between the lines (*Supplementary Materials*).

### 147 **Selection on Individuals’ Social Interaction Rules**

148 Polarization lines were significantly more cohesive than control lines (Fig. 2B; 3 mm, or 10%  
149 smaller median nearest neighbor distance; LMM:  $t = -5.5$ ,  $df = 170$ ,  $P < 0.001$ ), a finding that would  
150 not be expected if there were changes to individuals’ speeds but not in their social interactions (10).  
151 We therefore tested whether selection had altered the social interaction rules of the polarization  
152 compared to control lines. Many models and subsequent empirical work have identified that fish,  
153 including guppies, use attraction and alignment responses to coordinate their movements (5, 15). To  
154 test whether selection had changed the strength of these alignment or attraction rules, we first  
155 extracted the turning angles  $\alpha$  (Fig. 3B) that a fish made between its movement bursts. We then  
156 calculated the Spearman rank correlation over an entire trial between turning angles ( $\alpha$ ) and nearest  
157 neighbor directions ( $\beta$ ) to quantify attraction strength, and with neighbor orientations ( $\gamma$ ) to quantify  
158 alignment strength (*Supplementary Materials*). The strength of these correlations, therefore, acts as  
159 a proxy for the strength of these interactions. We found that fish from polarization lines had on  
160 average 23% higher correlations between turning angle and nearest neighbor orientation, and hence  
161 stronger alignment responses (LMM:  $t = 9.91$ ,  $df = 170$ ,  $P = 0.007$ ; Fig. 3D). There was also a non-

162 significant trend for polarization lines to have stronger attraction towards their nearest neighbor  
163 than control lines (LMM:  $t = 1.94$ ,  $df = 170$ ,  $P = 0.054$ ). When we included speed as a covariate in  
164 our models, fish from the polarization lines still showed 9% higher alignment strength than fish  
165 from control lines (LMM:  $t = 4.82$ ,  $df = 169$ ,  $P = 0.016$ ; see also the *Supplementary Materials*),  
166 showing that increased alignment responses were not only due to faster motion.

167 We then asked whether selection had changed the number of neighbors individuals were responding  
168 to during these attraction and alignment responses, using the centroid and mean orientation of the  $k$   
169 nearest neighbors to calculate  $\beta$  and  $\gamma$ , respectively (*Supplementary Materials*). The shape of the  
170 alignment and attraction responses, measured as a function of the number of influential neighbors,  
171 was qualitatively similar for both lines, declining after three to four neighbors in the case of  
172 alignment and plateauing at three to four neighbors in the case of attraction (Fig. 3D, 3E). This  
173 finding is reminiscent of the rule structuring in zonal models of collective motion, where alignment  
174 interactions occur with closer neighbors, and attraction responses with more distant neighbours  
175 (27). Although polarization-selected fish were not significantly more attracted to their nearest  
176 neighbor than control fish, attraction strength to multiple neighbors was stronger in the polarization  
177 than control lines. Attraction strength to  $k = 7$  nearest neighbors (i.e. the group centroid) increased  
178 in the polarization lines by 27% compared to control lines (LMM:  $t = 4.56$ ,  $df = 170$ ,  $P < 0.001$ ;  
179 Fig. 3E; see *Supplementary Materials* for male results). Selection may also have acted on the  
180 distance over which these alignment or attraction responses occurred. This is a common parameter  
181 in metric-based models of collective motion (21, 27), and we tested it by analysing occasions when  
182 the nearest neighbor was in front of a focal individual, and the focal individual either turned  
183 towards that neighbor with an attraction response ( $|\alpha - \beta| < 30$  degrees) or turned to align with that  
184 neighbor with an alignment response ( $|\alpha - \gamma| < 30$  degrees). We took the distance at which these  
185 responses occurred more frequently than by chance as a proxy for their interaction range  
186 (*Supplementary Materials*). We found no conclusive evidence that there were solid differences in  
187 either the attraction (LMM:  $t = 1.95$ ,  $df = 170$ ,  $P = 0.053$ ) or alignment ranges (LMM:  $t = -1.6$ ,  $df =$   
188  $170$ ,  $P = 0.11$ ) between the selection and control lines.

## 189 **Discussion**

190 Our results confirm that social interactions in a collective motion context are heritable, and that they  
191 can be rapidly shaped by directional artificial selection, leading to more polarized and cohesive  
192 groups. In particular, our selection regime changed three important aspects of individual behavior:  
193 1) speed, 2) the strength of the alignment response, and 3) the attraction strength to larger groups of  
194 conspecifics. Below we discuss the implications of these discoveries for our understanding of the  
195 interaction rules that lie behind evolutionary changes in collective motion.

196 First, increased speed has been suggested as an important and relatively simple mechanism behind  
197 more coordinated collective motion behavior (5, 15, 25). Importantly in our assays, the observed  
198 speed differences between the polarization and control lines were strongest in social contexts. For  
199 instance, the speed differences between lines were most prominent when close to conspecifics,  
200 suggesting that social facilitation may play an important role in how speed affects the increased  
201 polarization (28). Moreover, the observed differences in alignment were robust (albeit smaller) also  
202 when controlling for speed in the analysis, and we did not find any differences between polarization  
203 and control lines in our assays of physical swimming ability and behavioral stress responses. Hence,  
204 although our results indicate that speed changes play an important role for the behavioral  
205 differences between the polarization and control lines, we propose that such changes require a  
206 social context to be important for evolutionary shifts in collective motion.

207 After controlling for speed differences between the lines, fish from polarization lines were still  
208 more likely to align with their neighbors' directional heading than fish from control lines. The  
209 number of neighbors, or the range over which these alignment responses occurred, however, was  
210 not different between the selection lines. These results are consistent with how social  
211 responsiveness is often implemented in theoretical models of collective motion, where individuals  
212 weigh the tendency to travel in their own goal-orientated directions against the adoption of  
213 neighbors' directional headings (1, 29). It is possible, therefore, that selection acted on intrinsic  
214 differences in the social responsiveness of individuals, as has been predicted to exist in wild  
215 populations (14, 30). As well as increased alignment responses, polarization lines also showed  
216 stronger attraction responses to multiple conspecifics. While increased attraction is typically viewed  
217 in context of reducing predation risk through selfish herd effects (15, 31), increased attraction to  
218 others can also be viewed in the context of social decision-making, where individuals are often  
219 attracted towards larger numbers of neighbors (32). Our results suggest, therefore, that selection  
220 may have acted on how individuals weigh social information, ultimately leading to differences in  
221 group structure and social dynamics.

222 One possible explanation of the substantial change in polarization in only three generations is that  
223 the traits under selection have a simpler genetic background than is usually assumed. Examples do  
224 exist where seemingly complex behavior, such as burrowing behavior in mice, can have a relatively  
225 simple genetic architecture (33). In our experiment, however, the selection regime has changed  
226 several aspects of interaction rules in the polarization lines. We therefore view a very simple genetic  
227 background to these differences as unlikely, unless that architecture has pleiotropic effects across all  
228 these rules. Another explanation is due to the behaviors under selection here being a product of  
229 interactions between the behavior of a focal individual and the other individuals in the group. Such

230 social interactions have been suggested to be strongly influenced by indirect genetic effects, where  
231 expression of a trait in one individual alters expression of the trait across the social group, thereby  
232 amplifying the effect of selection (34). Indirect genetic effects could have played a role also in our  
233 experiment and increased the response to selection, but more work is needed to reveal the genetic  
234 architecture behind the observed differences.

235 In nature, which social rules evolve will ultimately depend on the selective forces present. Previous  
236 research has suggested that selective forces including the social environment (14), resource  
237 availability or distribution (8, 35), and predation risk (9, 15, 22) are likely to shape individuals'  
238 alignment and/or attraction responses. In turn, the social responses that evolve will have functional  
239 consequences for groups' abilities to track environmental gradients (36) and transfer information  
240 about detected threats or resources between group members (37, 38). However, to fully understand  
241 the evolution of these social rules, we also need to better understand the costs associated with  
242 evolving them. We found that increased coordination and cohesive behavior was associated with  
243 increased energy expenditure (i.e. increased speed). Similar energetic costs of coordination have  
244 been reported in flocks of birds (39). Future analyses on the polarization selection lines will  
245 investigate the costs and benefits of increased coordinated and collective movement in ecologically  
246 relevant settings.

247 It is noteworthy that the response to selection on polarization was weaker in males than in females.  
248 We specifically selected on female collective behavior in our experiment, and this could explain the  
249 weaker response in males. However, behaviors with strong fitness effects should have strong inter-  
250 sexual genetic correlations. The profound ecological differences between males and females in the  
251 guppy, with females having much higher propensity of shoaling (40), could explain the sex  
252 differences we observe. Our results certainly suggest that the genetic correlation between males and  
253 females for polarization behavior is relatively low, possibly due to differences in genetic  
254 architecture for social behavior between males and females (41).

255 In summary, our research has identified the social interaction rules that are affected by directional  
256 selection on polarization, and shown that such traits are susceptible to fast evolutionary changes. An  
257 integrated approach to understanding social behavior through artificial selection combined with  
258 detailed behavioral measurements now offers considerable opportunities to understand the evolution  
259 and maintenance of social decision-making and collective behavior.

## 260 **Material and Methods**

### 261 *Collective motion analysis*

262 Collective motion analyses were performed on 2565 mins of videos, obtained via a Point Grey  
263 Grasshopper 3 camera (FLIR Systems, resolution 2048 x 2048 px, frame rate 25 Hz), in MATLAB  
264 R2017b (details on video processing and data extraction can be found in (19)). Speed minima and  
265 maxima were found by first smoothing the speed profiles of individual fish (using a Savitzky-Golay  
266 filter degree three, span 12) and applying the *findpeaks* function. We found that turns in the  
267 trajectories typically came three frames (0.12 s) after a speed minimum, and accordingly applied  
268 this delay when calculating turning angles. For the assays of eight and single fish exploring the  
269 arenas, one of each measure was extracted per nine minute trial. The median distance from the edge  
270 of the arena was log-transformed and the polarization was transformed using an inverse logistic  
271 function.

### 272 *Swimming speed and boldness tests*

273 To test if selection changed the physical swimming ability of the fish, we measured the critical  
274 swimming speed of fish in a flow chamber (42) in 66 females and 62 males from the control and  
275 polarization selection lines. The flow chamber consisted of a 115 cm long transparent PVC pipe  
276 with an inside diameter of 1.8 cm through which aerated water was pumped at controllable speed  
277 ('swim tunnel'). We measured critical swimming speed by subjecting fish to increased velocity  
278 tests: the guppies were forced to swim against a current which was increased in discrete steps, until  
279 exhaustion occurred and they were swept against the outflow end of the tube. After a 2-minute  
280 acclimation period at a low velocity of 6.5 cm s<sup>-1</sup>, we increased the current velocity by 2.2 cm s<sup>-1</sup>  
281 every 30 s until the guppy reached exhaustion and was unable to detach itself from the outflow  
282 mesh for 3 seconds. Temperature was maintained at 25.0°C ± 1.5°C. Results are shown in Fig. S2  
283 and Table S1.

284 We also measured boldness and exploration of 60 females and 60 males in a standard emergence  
285 test; in a 50 l tank with 3 cm of water. The starting compartment (20 x 10 cm) was separated from  
286 the exploration compartment (20 x 40 cm) by an opaque partition with an eight cm wide opening.  
287 After two minutes of acclimation in the starting compartment an opaque trap door was lifted to  
288 allow access to the exploration compartment. The time it took until individuals left the starting  
289 compartment was used as indicator of boldness and the number of 5 x 5 cm plots visited (15 in  
290 total, every time a new plot was visited, this was added to the total area visited) was used as  
291 indicator for exploratory tendencies 3. Non-emerged fish after the maximum time of 10 minutes  
292 were removed from the analysis (15 from each set of female lines, two from male polarization lines,  
293 four from male control lines). The time to exit a shelter was used as a measure of a fish's boldness,  
294 and the area explored by each fish was used as a measure of their exploratory tendencies. Both  
295 measures were log-transformed. Results are shown in Fig. S3 and Table S1.

296 *Statistics*

297 We tested for differences between selection lines using linear mixed-effect models. Separate models  
298 were used for individual trials and groups of eight, as well as for males and females. Selection line  
299 was incorporated as a fixed effect. For tracked motion assays, mean body size (estimated from  
300 IDTracker) was incorporated as a covariate, as well as median speed when controlling for activity.  
301 Replicate was used as a random effect for the intercept and the selection effect. Normality of  
302 residuals were checked using Kolmogorov-Smirnov tests, with a maximum KS statistic of 0.107.  
303 Residuals were plotted against fitted values to visually check for correlations and heteroscedasticity.  
304 Analyses were done in MATLAB R2017b.

305

306

307 **References**

308

- 309 1. I. D. Couzin, J. Krause, N. R. Franks, S. A. Levin, Effective leadership and decision-making in animal groups  
310 on the move. *Nature* **433**, 513-516 (2005).
- 311 2. C. Breder Jr, Equations descriptive of fish schools and other animal aggregations. *Ecology* **35**, 361-370 (1954).
- 312 3. U. Lopez, J. Gautrais, I. D. Couzin, G. Theraulaz, From behavioural analyses to models of collective motion in  
313 fish schools. *Interface focus* **2**, 693-707 (2012).
- 314 4. D. S. Calovi *et al.*, Swarming, schooling, milling: phase diagram of a data-driven fish school model. *New*  
315 *Journal of Physics* **16**, 015026 (2014).
- 316 5. J. Gautrais *et al.*, Deciphering interactions in moving animal groups. *PLoS Comput Biol* **8**, e1002678 (2012).
- 317 6. B. Pettit, A. Perna, D. Biro, D. J. Sumpter, Interaction rules underlying group decisions in homing pigeons.  
318 *Journal of The Royal Society Interface* **10**, 20130529 (2013).
- 319 7. A. Strandburg-Peshkin *et al.*, Visual sensory networks and effective information transfer in animal groups.  
320 *Current Biology* **23**, R709-R711 (2013).
- 321 8. A. M. Hein *et al.*, The evolution of distributed sensing and collective computation in animal populations. *Elife*  
322 **4**, e10955 (2015).
- 323 9. C. C. Ioannou, V. Guttal, I. D. Couzin, Predatory fish select for coordinated collective motion in virtual prey.  
324 *Science* **337**, 1212-1215 (2012).
- 325 10. J. W. Jolles, N. J. Boogert, V. H. Sridhar, I. D. Couzin, A. Manica, Consistent Individual Differences Drive  
326 Collective Behavior and Group Functioning of Schooling Fish. *Current Biology* **27**, 2862-2868. e2867 (2017).
- 327 11. L. M. Aplin *et al.*, Consistent individual differences in the social phenotypes of wild great tits, *Parus major*.  
328 *Animal Behaviour* **108**, 117-127 (2015).
- 329 12. J. E. Herbert-Read, Understanding how animal groups achieve coordinated movement. *Journal of*  
330 *Experimental Biology* **219**, 2971-2983 (2016).
- 331 13. A. K. Greenwood, A. R. Wark, K. Yoshida, C. L. Peichel, Genetic and neural modularity underlie the  
332 evolution of schooling behavior in threespine sticklebacks. *Current Biology* **23**, 1884-1888 (2013).
- 333 14. M. Wolf, F. J. Weissing, An explanatory framework for adaptive personality differences. *Philosophical*  
334 *Transactions of the Royal Society B-Biological Sciences* **365**, 3959-3968 (2010).
- 335 15. J. E. Herbert-Read *et al.*, How predation shapes the social interaction rules of shoaling fish. *Proceedings of the*  
336 *Royal Society B-Biological Sciences* **284**, 20171126 (2017).
- 337 16. A. E. Magurran, *Evolutionary Ecology: the trinidadian guppy.*, (Oxford University Press, Oxford, 2005).
- 338 17. D. P. Croft *et al.*, Sex-biased movement in the guppy (*Poecilia reticulata*). *Oecologia* **137**, 62-68 (2003).
- 339 18. A. Szorkovszky *et al.*, An efficient method for sorting and quantifying individual social traits based on group  
340 level behaviour. *Methods in Ecology and Evolution* **8**, 1735-1744 (2017).
- 341 19. A. Szorkovszky *et al.*, Assortative interactions revealed by sorting of animal groups. *Animal Behaviour* **142**,  
342 165-179 (2018).
- 343 20. T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, O. Shochet, Novel type of phase transition in a system of self-  
344 driven particles. *Physical review letters* **75**, 1226 (1995).
- 345 21. D. Strömbom, Collective motion from local attraction. *Journal of theoretical biology* **283**, 145-151 (2011).
- 346 22. A. J. Wood, G. J. Ackland, Evolving the selfish herd: emergence of distinct aggregating strategies in an

- 347 individual-based model. *Proceedings of the Royal Society of London B: Biological Sciences* **274**, 1637-1642  
348 (2007).
- 349 23. M. Ballerini *et al.*, Interaction ruling animal collective behavior depends on topological rather than metric  
350 distance: Evidence from a field study. *Proceedings of the national academy of sciences* **105**, 1232-1237  
351 (2008).
- 352 24. N. W. Bode, D. W. Franks, A. J. Wood, Making noise: emergent stochasticity in collective motion. *Journal of*  
353 *theoretical biology* **267**, 292-299 (2010).
- 354 25. S. Mishra, K. Tunström, I. D. Couzin, C. Huepe, Collective dynamics of self-propelled particles with variable  
355 speed. *Physical Review E* **86**, 011901 (2012).
- 356 26. A. Pérez-Escudero, J. Vicente-Page, R. C. Hinz, S. Arganda, G. G. De Polavieja, idTracker: tracking  
357 individuals in a group by automatic identification of unmarked animals. *Nature methods* **11**, 743-748 (2014).
- 358 27. I. D. Couzin, J. Krause, R. James, G. D. Ruxton, N. R. Franks, Collective memory and spatial sorting in animal  
359 groups. *Journal of theoretical biology* **218**, 1-11 (2002).
- 360 28. M. M. Webster, A. J. Ward, Personality and social context. *Biological Reviews* **86**, 759-773 (2011).
- 361 29. I. D. Couzin *et al.*, Uninformed individuals promote democratic consensus in animal groups. *science* **334**,  
362 1578-1580 (2011).
- 363 30. S. R. Dall, A. I. Houston, J. M. McNamara, The behavioural ecology of personality: consistent individual  
364 differences from an adaptive perspective. *Ecology letters* **7**, 734-739 (2004).
- 365 31. W. D. Hamilton, Geometry for the selfish herd. *Journal of theoretical Biology* **31**, 295-311 (1971).
- 366 32. A. J. Ward, D. J. Sumpter, I. D. Couzin, P. J. Hart, J. Krause, Quorum decision-making facilitates information  
367 transfer in fish shoals. *Proceedings of the National Academy of Sciences* **105**, 6948-6953 (2008).
- 368 33. J. Weber, B. Peterson, H. Hoekstra, Discrete genetic modules are responsible for the evolution of complex  
369 burrowing behaviour in deer mice. *Nature* **493**, 4202-4405 (2013).
- 370 34. F. Santostefano, A. J. Wilson, P. T. Niemelä, N. J. Dingemanse, Indirect genetic effects: a key component of  
371 the genetic architecture of behaviour. *Scientific reports* **7**, 10235 (2017).
- 372 35. V. Guttal, I. D. Couzin, Social interactions, information use, and the evolution of collective migration.  
373 *Proceedings of the national academy of sciences* **107**, 16172-16177 (2010).
- 374 36. A. Berdahl, C. J. Torney, C. C. Ioannou, J. J. Faria, I. D. Couzin, Emergent sensing of complex environments  
375 by mobile animal groups. *Science* **339**, 574-576 (2013).
- 376 37. S. B. Rosenthal, C. R. Twomey, A. T. Hartnett, H. S. Wu, I. D. Couzin, Revealing the hidden networks of  
377 interaction in mobile animal groups allows prediction of complex behavioral contagion. *Proceedings of the*  
378 *National Academy of Sciences* **112**, 4690-4695 (2015).
- 379 38. A. Strandburg-Peshkin, D. R. Farine, I. D. Couzin, M. C. Crofoot, Shared decision-making drives collective  
380 movement in wild baboons. *Science* **348**, 1358-1361 (2015).
- 381 39. L. Taylor *et al.*, Birds invest wingbeats to keep a steady head and reap the ultimate benefits of flying together.  
382 *PLoS Biology*, (2019).
- 383 40. D. P. Croft *et al.*, Assortative interactions and social networks in fish. *Oecologia* **143**, 211-219 (2005).
- 384 41. J. E. Mank, The transcriptional architecture of phenotypic dimorphism. *Nature Ecology & Evolution* **1**, 0006  
385 (2017).
- 386 42. A. Kotrschal *et al.*, Artificial selection on relative brain size reveals a positive genetic correlation between  
387 brain size and proactive personality in the guppy. *Evolution* **68**, 1139-1149 (2014).
- 388

## 389 **Acknowledgements**

390 **General:** We thank Anna Rennie, Eduardo Trejo and Annika Boussard for help with fish  
391 husbandry.

392 **Funding:** This work was supported by the Knut and Alice Wallenberg Foundation (102 2013.0072  
393 to DS, NK and KP) and the Swedish Research Council (2016-03435 to NK, 2017-04957 to AK,  
394 2018-04076 to JHR)

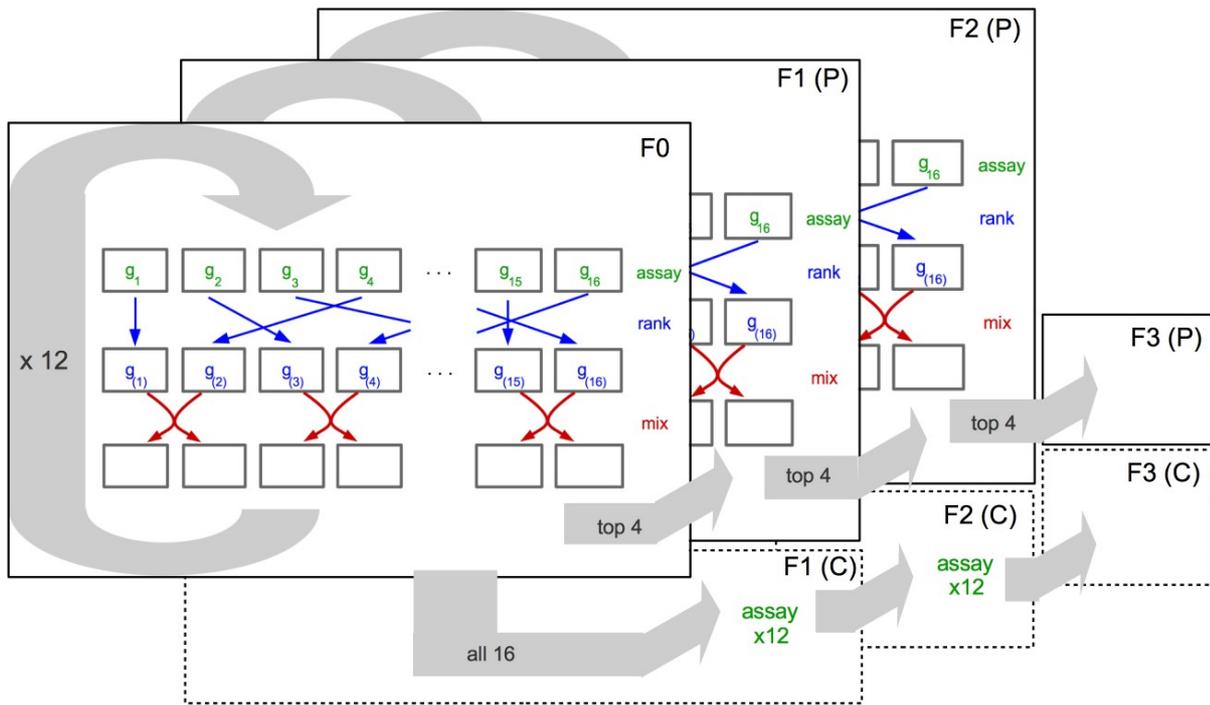
395 **Author contributions:** NK, KP, and DS conceived the idea to the study. AK, AS, MR, SDB, JHR,  
396 HZ, KP DS, and NK designed the details of the selection procedure. MR set up the filming  
397 apparatus, AK performed and AS analyzed the data of the selection experiment. AK, AF, LSA,

398 ALF, and GB performed and AK, AS and SDB analyzed the data of the behavioural assays. AK,  
399 AS and JHR created the figures. AK, AS, JHR, JEM, NIB, DS, and NK wrote the manuscript. All  
400 authors contributed to the final version of the manuscript.

401 **Competing interests:** No competing interests are declared

402 **Data availability:** All data and code will be deposited in Dryad.

403

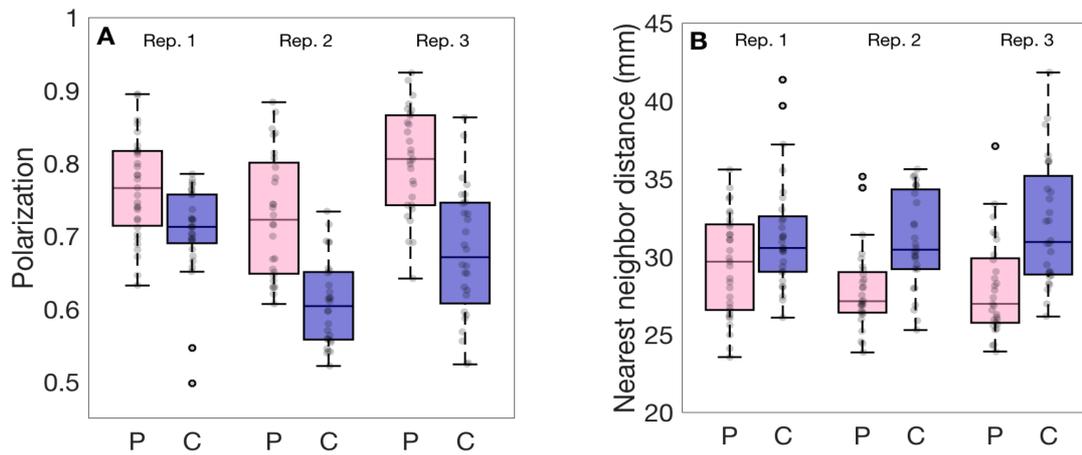


404

405

406 **Fig. 1. Schematic of one of three independent replicates of the selection experiment.** Each  
 407 layer represents a generation of females. Arrows within a layer illustrate the sorting procedure  
 408 where we identified fish that formed the most polarized groups. To do this, groups were first  
 409 assayed for their group polarization. Here variables  $g_1$  to  $g_{16}$  denote the 16 groups'  
 410 polarization scores in round  $t$ . These scores were subsequently ranked (blue arrows) with  $g_{(1)}$  to  
 411  $g_{(16)}$  denoting the ranked scores from lowest to highest. Following this, half of the group  
 412 members were mixed with adjacently ranked groups (red arrows). This ranking and sorting  
 413 procedure was repeated 12 times (circular grey arrow) before 26 fish from the top four ranked  
 414 groups were bred for the polarization lines, and 26 fish from remaining 16 groups were bred for  
 415 the control lines. This sorting procedure was repeated three times for the polarization lines  
 416 (indicated by the layers), whereas fish from the control group experienced the same assaying and  
 417 sorting, except fish from these control lines were not ranked.

418



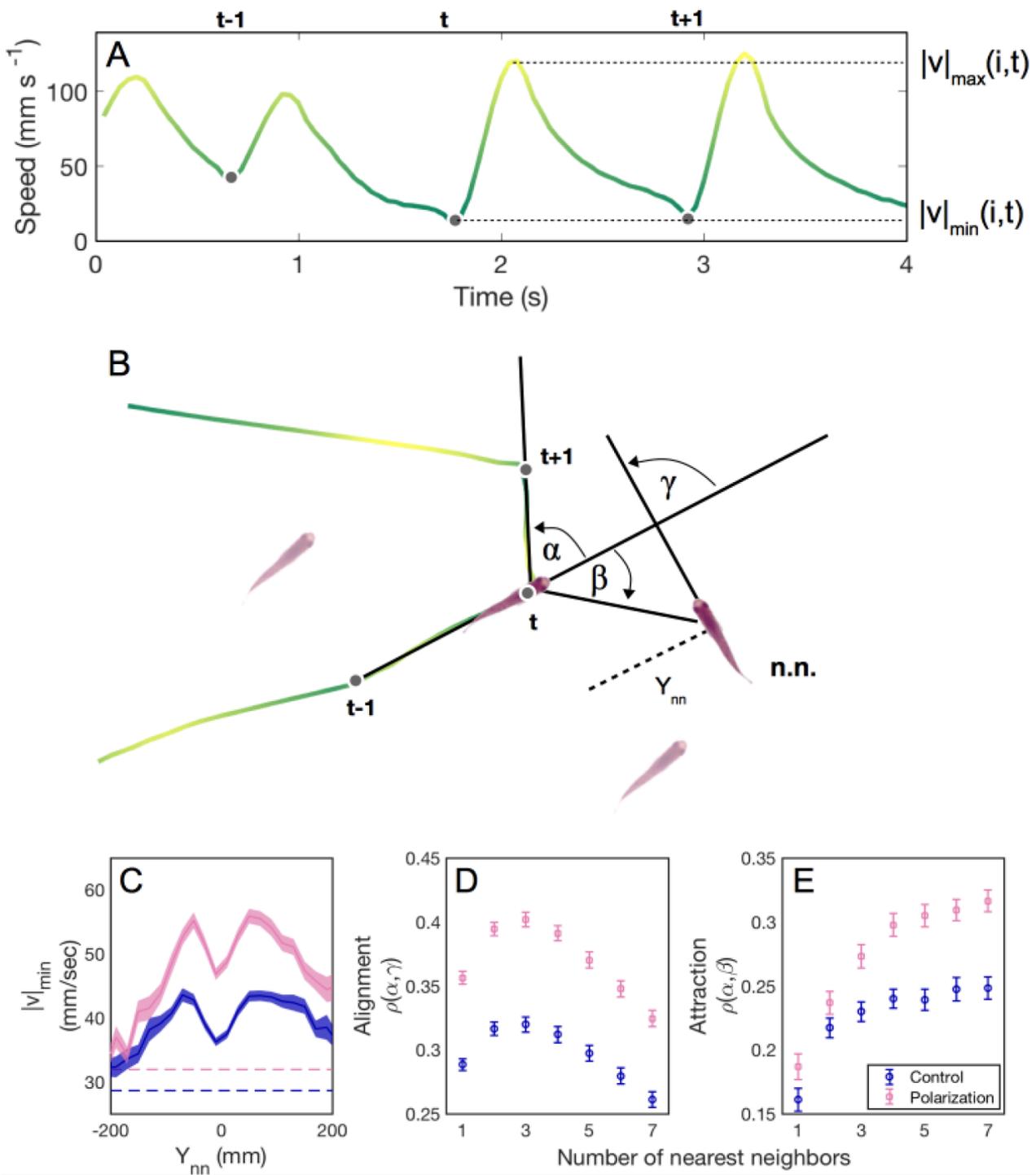
419

420

421 **Fig. 2. Polarization and nearest neighbor distance in groups of guppies artificially selected for**  
422 **polarization.** Boxplots of (A) median polarization and (B) median nearest neighbor distance for  
423 groups of eight females in polarization selected (pink boxed) or control lines (blue boxes). Replicate  
424 lines 1, 2 and 3 are denoted above the boxes. Grey markers show individual data points (i.e. trials).  
425 Horizontal lines indicate medians, boxes indicate the interquartile range, and whiskers indicate all  
426 points within 1.5 times the interquartile range.

427

428



429

430

431 **Fig. 3. Burst and glide analysis and inferring social interactions of guppies artificially selected**  
 432 **for polarization.** (a) Time series containing three consecutive speed minima (dots) followed by  
 433 bursting events. (b) The corresponding trajectory for fish  $i$  (in the center). The positions at the  
 434 preceding and following speed minima are used to calculate the turning angle  $\alpha$  of fish  $i$  at time  $t$ . In  
 435 this example, the turn has the same sign as the nearest neighbor orientation (i.e. alignment)  $\gamma$ , and  
 436 the opposite sign to the attraction angle  $\beta$ . (c-e) The social interactions of females in the control

437 lines (blue) and polarization selection lines (pink) in response to nearest neighbors. (c) The mean  
438 speed minimum when the nearest neighbor is in front (+) or behind (−) by a distance  $Y_{nm}$ . The error  
439 region shows the standard error in the mean over trials. (d) Alignment and (e) attraction responses  
440 to the geometric center of  $k$  nearest neighbors, where  $k$  ranges from 1 (nearest neighbor) to 7 (all  
441 conspecifics). The Spearman correlations  $\rho$  were computed for each  $k$  and for each trial for all  $\beta$  and  
442  $\gamma$  with absolute values of less than 90 degrees. The set of  $\beta$  was additionally restricted to time points  
443 where the  $k$  neighbors were all less than 200 mm from the focal fish (*Supplementary Materials*).  
444 Means (symbols) and standard errors (bars) were calculated for each selection line from these  
445 correlation coefficients.

446

## **Supplementary Information for**

### **Rapid evolution of coordinated and collective movement in response to artificial selection**

### **Evolution of social interaction rules**

Alexander Kotrschal<sup>\*1,2†</sup>, Alexander Szorkovszky<sup>1,3†</sup>, James Herbert-Read<sup>4,5</sup>,

Natasha I. Bloch<sup>6</sup>, Maksym Romenskyy<sup>7</sup>, Séverine Denise Buechel<sup>1</sup>, Ada

Fontrodona Eslava<sup>1,8</sup>, Laura Sánchez Alòs<sup>1</sup>, Hongli Zeng<sup>9</sup>, Audrey Le Foll<sup>1</sup>, Ganaël

Braux<sup>1</sup>, Kristiaan Pelckmans<sup>10</sup>, Judith E. Mank<sup>11,12</sup>, David Sumpter<sup>3</sup> & Niclas Kolm<sup>1</sup>

† equal contribution

\* corresponding author: [alexander.kotrschal@zoologi.su.se](mailto:alexander.kotrschal@zoologi.su.se)

## Artificial Selection

Fig. S1 shows the polarization of the original F0 female groups, and the selection and control lines in subsequent generations.

## Additional collective motion analysis

**Complimentary burst and glide results.** Female groups from the polarization lines made more frequent bursts than the corresponding control lines (mean time between bursts was shorter by 0.022 s or 2.7%; LMM:  $t = 3.01$ ,  $df = 170$ ,  $P = 0.003$ ) and made smaller turns (0.042 radians or 6.9%; LMM:  $t = 5.79$ ,  $df = 170$ ,  $P < 0.001$ ). This was most likely due to the negatively correlated biomechanics of speed and turning combined with the arena's geometry.

**Thigmotaxis analysis.** To control for potential differences in thigmotaxis (attraction to the walls of the arena), that could affect the potential for polarization if one group tended to swim nearer the edges, we analyzed the median distances from the edge of the arena across the selection lines. The polarization line females were on average 2.6 mm further from the center of the arena in trials with single fish (LMM:  $t = 2.12$ ,  $df = 117$ ,  $P = 0.036$ ) compared to the single fish controls (257.9 mm), and 3.1 mm further from the center of the arena in groups (LMM:  $t = 3.49$ ,  $df = 170$ ,  $P < 0.001$ ) compared to the control groups (253.9 mm). We therefore repeated our analysis only including frames from the videos where the mean distance from the arena edge was more than 50 mm (i.e. less than 225 mm from the center). This analysis showed that polarization line females were still 10.2% more aligned than the control lines (LMM:  $t = 4.34$ ,  $df = 170$ ,  $P < 0.001$ ), demonstrating that differences in thigmotaxis were not driving the differences in polarization between the polarization and control lines.

**Conspecific speed.** Attraction and alignment turning responses to the nearest neighbor (quantified by the Spearman rank correlation, see Fig. S4) varied with the neighbor's speed and position. To illustrate this, we plotted aggregated attraction and alignment responses as a function of both forward distance to nearest neighbor and speed of this neighbor at the bursting time of the focal fish (Fig. S5, S6). Attraction responses were strongest when the neighbor was close in front, while alignment responses were strongest when the neighbor was close in front and travelling more quickly. By subtracting the control line heat maps from the polarization line heat maps (Fig. S5 C, F), it is clear that the female fish from the polarization lines show an increased response in these regions. That is, for comparable speeds and positions, the polarization line females show an increased correlation between turning response and the nearest neighbor's position and orientation, lending further support that social interactions in the females differed between selection lines. This was less apparent in the males (Fig. S6 C, F), although these

patterns occurred in the same direction as the females.

**Multiple neighbor responses.** We used both metric and topological models to test turning responses to multiple neighbors. The topological attraction model used the mean position of the nearest  $k$  neighbors as the  $X$  variable for the Spearman correlation. See Fig. S4 for a typical trial for  $k = 1$  and  $k = 7$ . In all per-trial correlations, only data points with  $X < 90$  degrees were used, as the dependence is monotonic in this region. The topological attraction correlations for females as a function of  $k$  are shown in Fig. S7 A.

In the metric attraction model, the  $X$  variable is the mean position of all neighbors within a distance  $r$ , rather than a specific number of neighbors. As can be seen in Fig. S7 B, this correlation reached a peak around 200 mm and then declined. This weakening of the response to distant neighbors is consistent with the decline in the topological attraction model for large  $k$ . Restricting the topological model to neighbors within 200 mm hence partially controls for differences in cohesion between lines. In this case, the topological attraction reached a plateau at  $k > 3$  for both lines (see main text Fig. 3 D, E).

In the topological and metric alignment models (Fig. S7 C, D), the  $X$  variable is the mean orientation of the neighbors within the first  $k$  neighbors or radius  $r$  respectively. The orientation of the focal fish was not included in this calculation.

**Range calculations.** We calculated the range of a response ( $A$ ) using kernel-smoothed distributions of nearest neighbor distances,  $R$ . A typical response range for  $A$  can be defined as a range of  $r$  satisfying

$$P(A|R = r) > P(A) \quad [1]$$

which by Bayes' rule is equivalent to

$$P(R = r|A) > P(R = r) \quad [2]$$

Hence, for each trial, we calculated these two probability density functions of the nearest neighbor distance,  $R$ . The density function corresponding to  $P(R = r)$  is calculated using all decision points in which the neighbor is in front of the focal fish. All nearest neighbor distances at these points are input to a kernel-smoothed density estimator with bandwidth of 10 mm. The density function  $P(R = r|A)$  is calculated in the same way, but with the extra restriction that there is a response  $A$  (i.e. the turning angle is within 30 degrees of the vector corresponding to the neighbor position (attraction) or orientation (alignment)). We then modify the above expression slightly to define the range  $r_{\max}$  as:

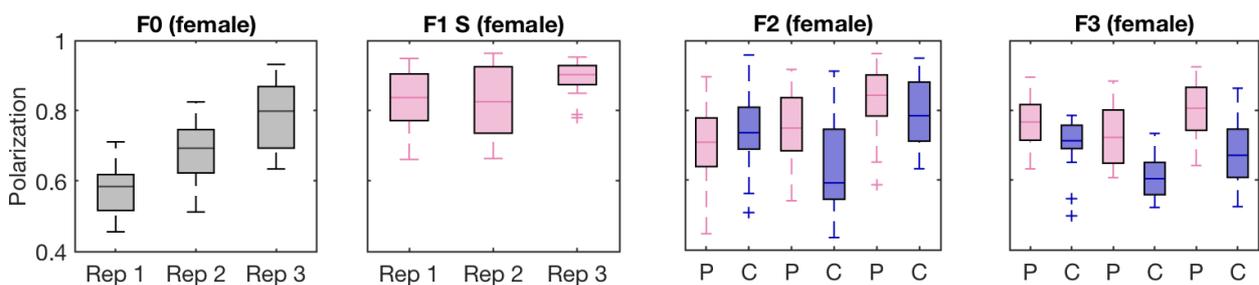
$$r_{\max} = \sup[r : P(R = r|A) > P(R = r) + \epsilon] \quad [3]$$

We used a small constant offset  $\epsilon = 2.5 \times 10^{-4}$  to account for small fluctuations in the tails of the distributions (see Fig. S10).

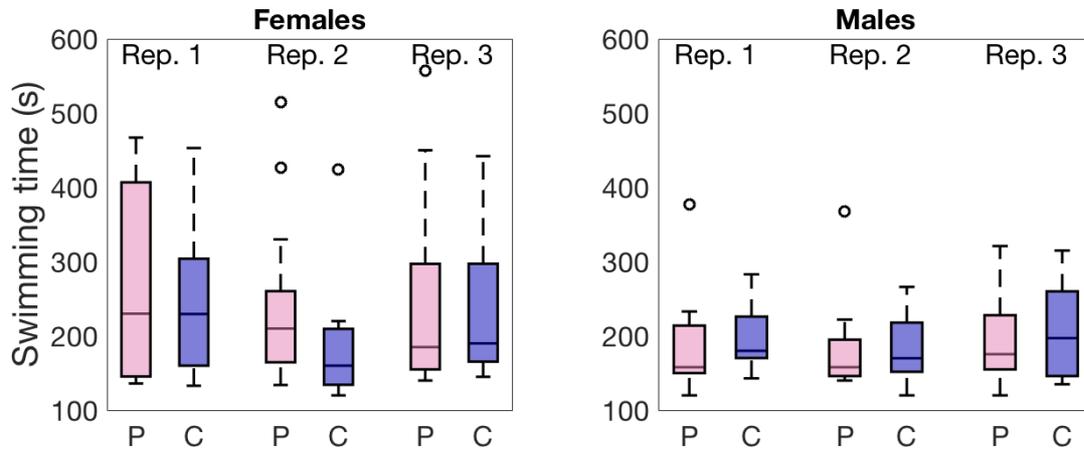
11

## Male analysis

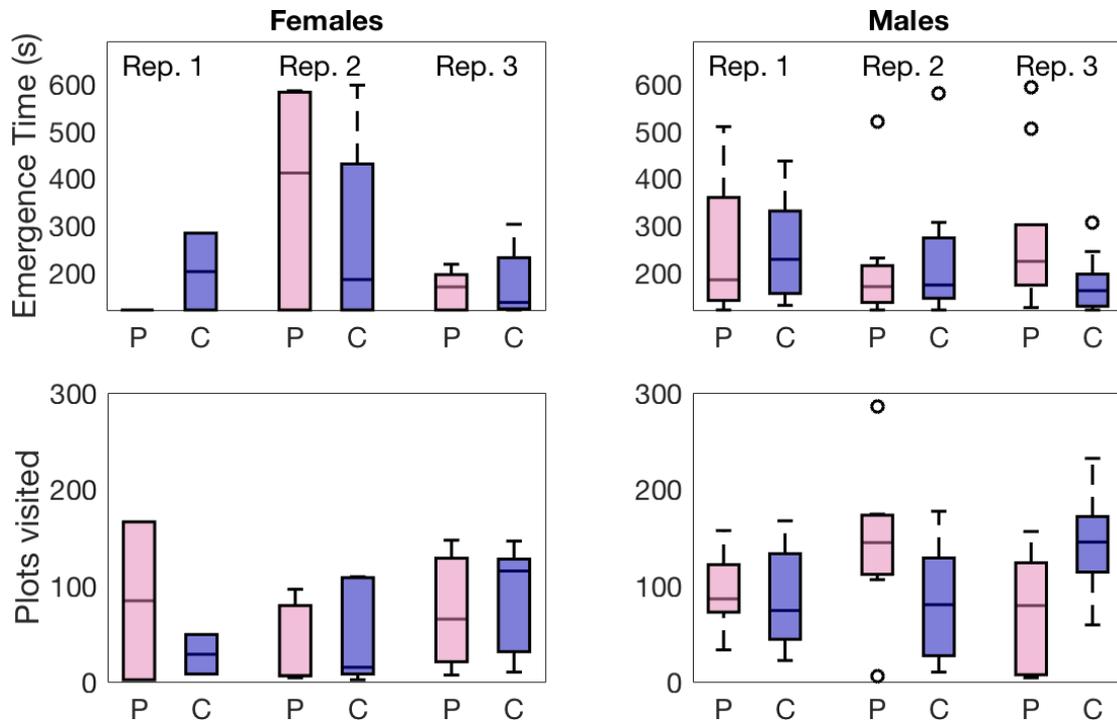
When testing offspring from the third generation of selection, we did not find any significant differences in the polarization of male groups between the polarization ( $n = 56$ ) and control lines ( $n = 56$ ), although trends were in the same direction as the females (LMM:  $t = 0.944$ ,  $df = 109$ ,  $P = 0.26$ ). While males did not significantly differ in polarization between the lines, in concordance with the females' results, males' median speed was higher by  $5.7 \text{ mm s}^{-1}$  (9%) in polarization compared to the control lines when tested in groups (LMM:  $t = 2.95$ ,  $df = 109$ ,  $P = 0.004$ ), but not when tested alone (LMM:  $t = 0.708$ ,  $df = 109$ ,  $P = 0.48$ ). The nearest neighbor alignment responses and the group attraction responses were respectively 10% and 22% stronger in males from the polarization lines compared to control lines (nearest neighbor alignment LMM:  $t = 2.76$ ,  $df = 109$ ,  $P = 0.007$ ; group attraction LMM:  $t = 3.21$ ,  $df = 109$ ,  $P = 0.002$ ; Fig. S9), again, a finding that was consistent with results from the females. Together, these results suggest that although males from polarization and control lines differed in speed and social interactions in the same way as female fish, this did not generate the same strong differences in polarization that were observed in the females. This may be due to the generally reduced social tendencies of males compared to females, which could be due to differences in social behavior that are sexually linked.



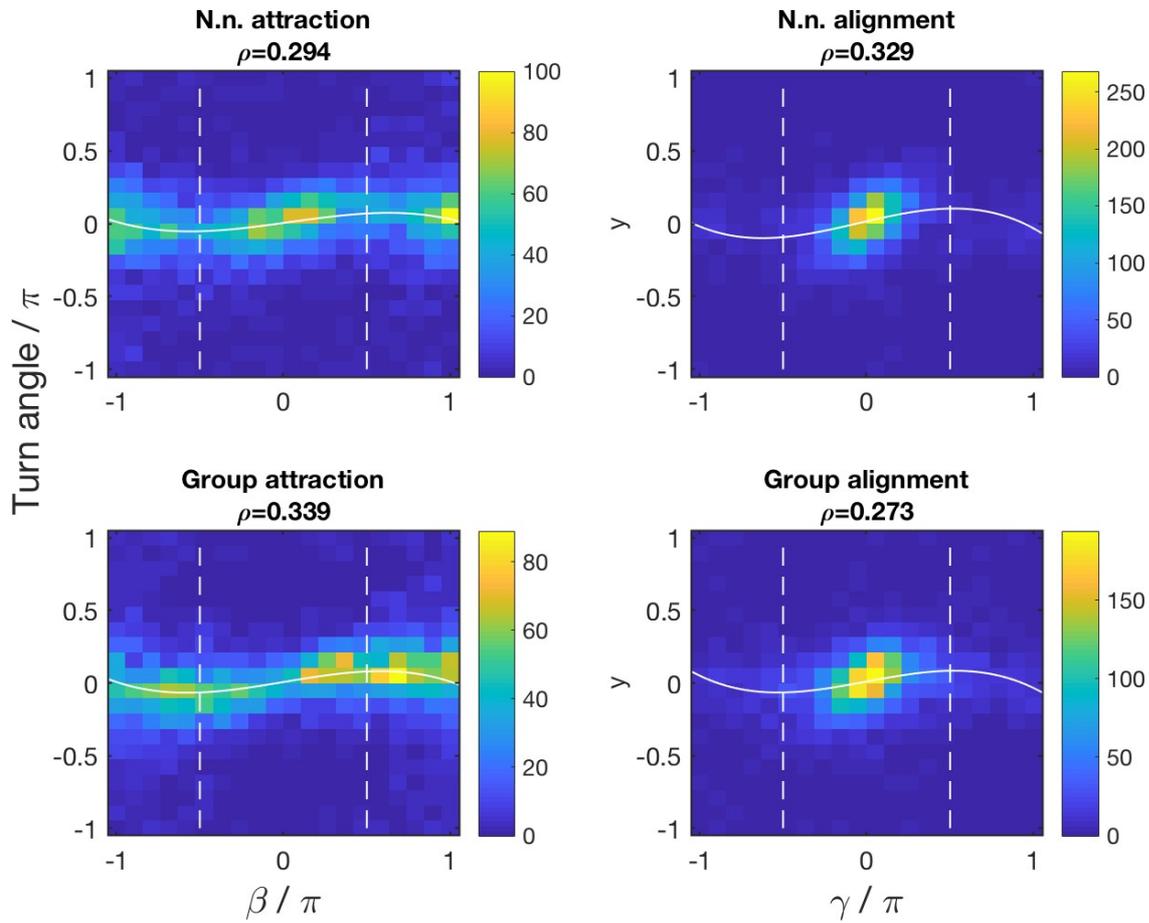
**Fig. S1.** Polarization of females in F0 (base population), F1 (first sorting round, control lines not filmed), F2 and F3. Higher polarization values indicate more coordinated schooling behavior. Horizontal lines indicate medians, boxes indicate the interquartile range, and whiskers indicate all points within 1.5 times the interquartile range. P: Polarization line (pink bars), C: Control line (blue bars).



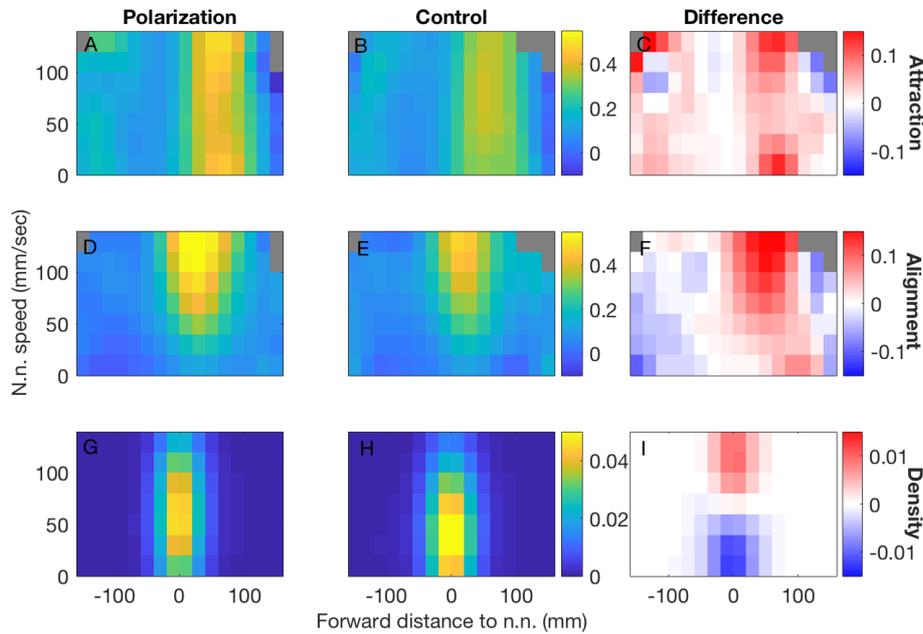
**Fig. S2.** In a swim tunnel with increasing laminar water current polarization (P, pink) and control (C, blue) lines showed similar maximal swimming speeds in females (left panel) and males (right panel). See Table S1 for statistics on log-transformed variables. Note that swimming time and speed are equivalent as swimming speed was increased at constant time intervals. Horizontal lines indicate medians, boxes indicate the interquartile range, and whiskers indicate all points within 1.5 times the interquartile range.



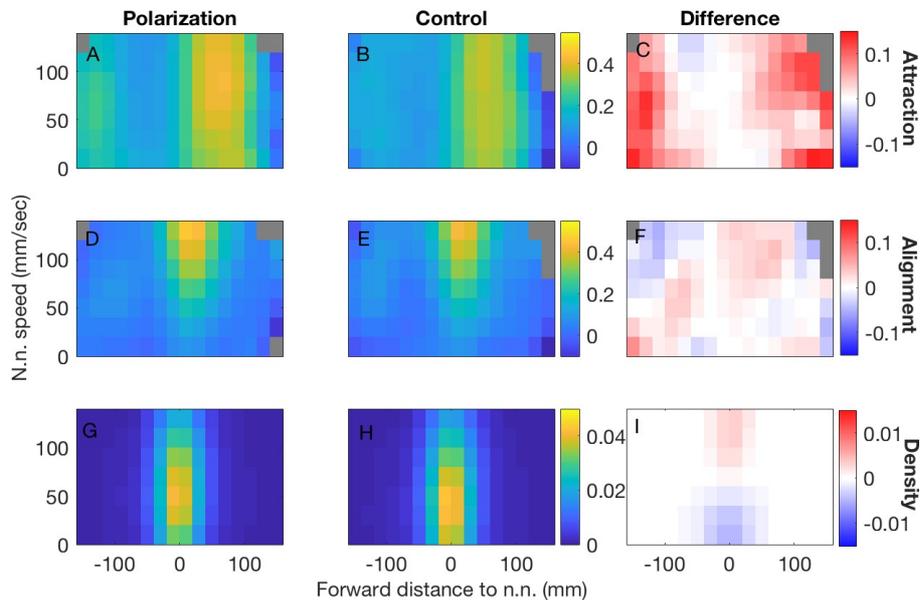
**Fig. S3.** An emergence test revealed no differences in emergence time ('boldness' top panel) or the number of plots visited ('exploration' bottom panel) between polarization (P, pink) and control (C, blue) lines in females (left panels) or males (right panels). After log-transformation, no differences were found in females or males (see Table S1). Horizontal lines indicate medians, boxes indicate the interquartile range, and whiskers indicate all points within 1.5 times the interquartile range.



**Fig. S4.** Rank correlations for a single female trial. Each panel is a bivariate histogram of the turning angle (vertical axis) against a predictor (horizontal axis). The color bar shows the number of data points in each bin. A third-order polynomial is fit to each set of points and is shown as a solid white line. The Spearman correlations shown above each panel are calculated from all points where the absolute predictor angle is less than 90 degrees (i.e. within the dotted white lines).

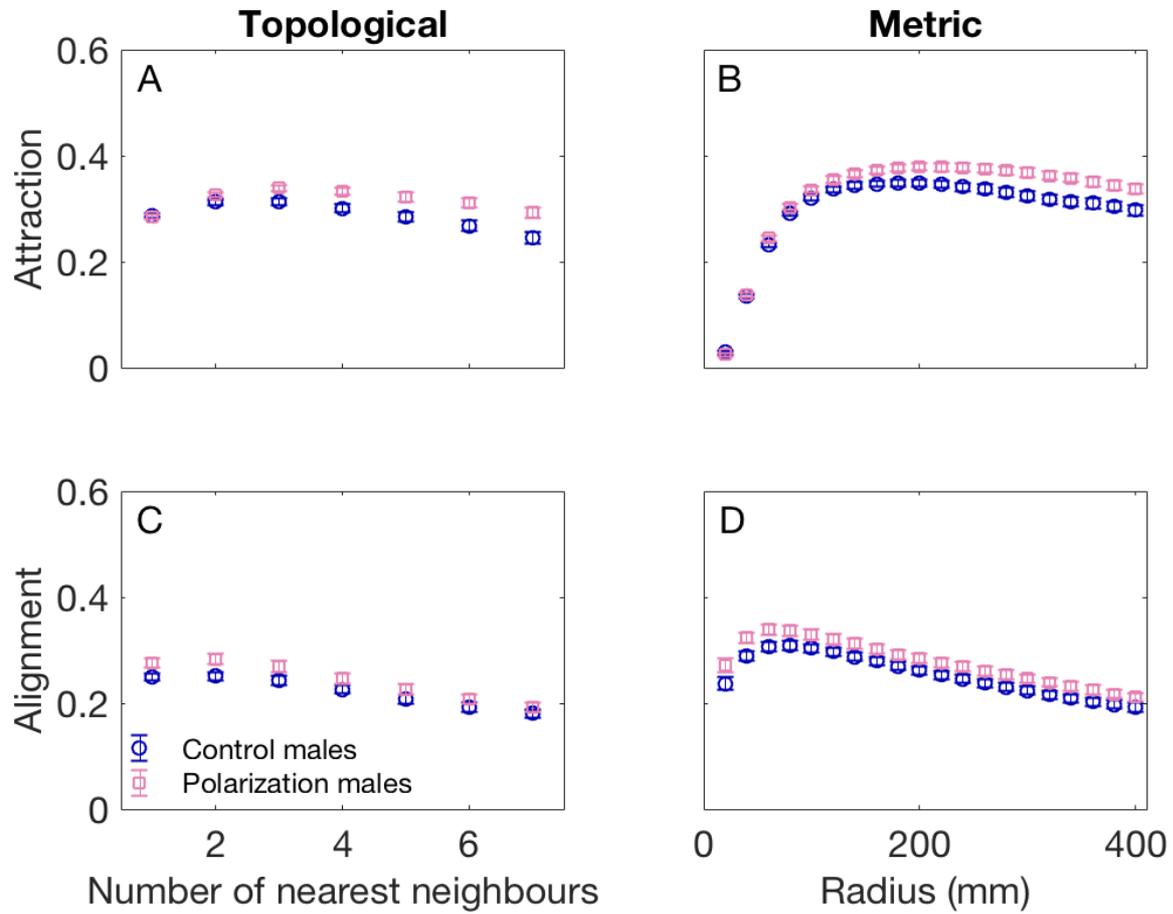


**Fig. S5.** Typical burst and glide behavior of females as a function of nearest neighbor position and speed. In all panels the horizontal axis is the forward distance to the nearest neighbor and the vertical axis is the nearest neighbor speed. In each bin within panels A, B, D, and E, the Spearman rank correlation was calculated between the turning angles and the vector corresponding to the position or orientation of the nearest neighbor to quantify the attraction or alignment strengths respectively. Panels C&F show differences in rank correlation between the two panels to the left. Panels G&H show the density of data points, while panel I shows the difference in densities.

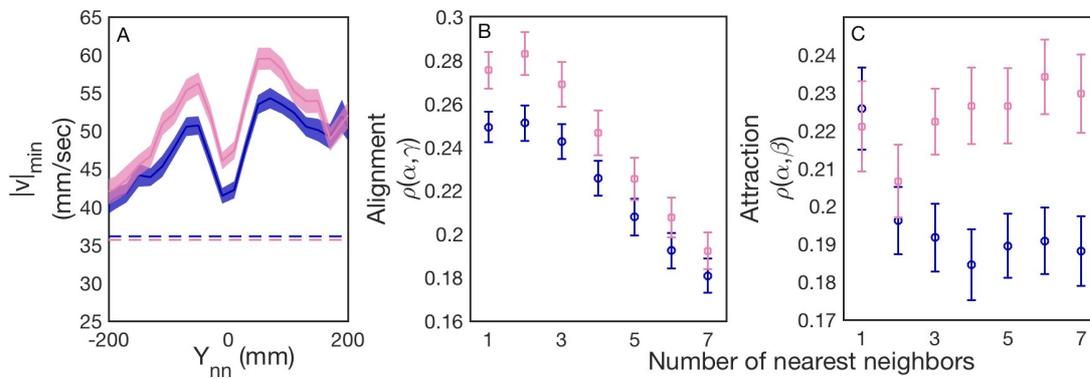


**Fig. S6.** Typical burst and glide behavior of males as a function of nearest neighbor position and speed. All panels as in Fig S4. In all panels the horizontal axis is the forward distance to the nearest neighbor and the vertical axis is the nearest neighbor speed. In each bin within panels A, B, D, and E, the Spearman rank correlation was calculated between the turning angles and the vector corresponding to the position or orientation of the nearest neighbor to quantify the attraction or alignment strengths respectively. Panels C and F show differences in rank correlation between the two panels to the left. Panels G and H show the density of data points, while panel I shows the difference in densities.

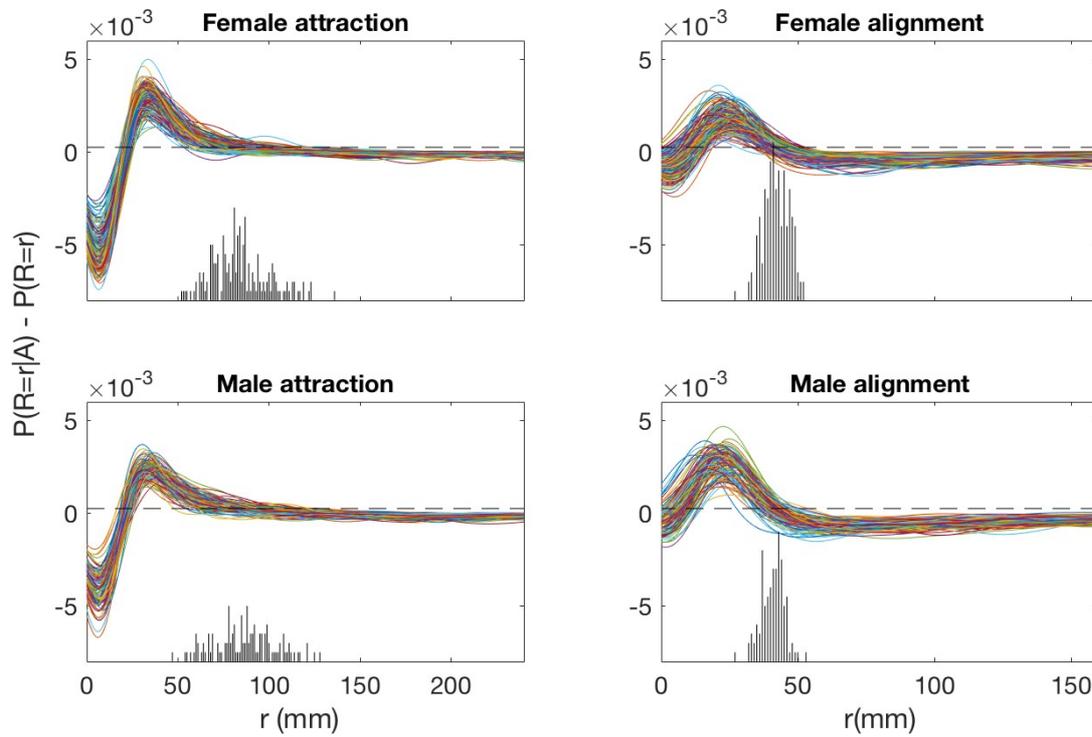




**Fig. S8.** Male alignment and attraction responses vs. topological distance (A, C) and metric distance (B, D). Shown are means of the per-trial rank correlations for the selection and control lines; with error bars indicating the standard errors over trials.



**Fig. S9.** Typical burst and glide behavior of males in the control lines (blue) and polarization selection lines (pink) in response to nearest neighbors. (A) Mean speed minimum when the nearest neighbor is in front (+) or behind (−) by a distance  $Y_{nn}$ . The error region shows the standard error in the mean over trials. (B) Alignment and (C) Attraction responses to the geometric center of  $k$  nearest neighbors, where  $k$  ranges from 1 (nearest neighbor) to 7 (all conspecifics). The Spearman correlations  $\rho$  were computed for each  $k$  and for each trial for all  $\beta$  and  $\gamma$  with absolute values of less than 90 degrees. The set of  $\beta$  was additionally restricted to time points where the  $k$  neighbors were all less than 200 mm from the focal fish. Means (symbols) and standard errors (bars) were calculated for each selection line from these correlation coefficients.



**Fig. S10.** Calculation of attraction and alignment ranges. In each panel, each curve corresponds to the difference in kernel-smoothed distributions for one trial of eight fish. A region where the curve is above zero indicates a response above the average level. The dotted line shows the cut off at  $s$ . The histograms below show the distributions of the final points in  $r$  above the cutoff, which are used as the attraction and alignment ranges.

**Table S1. Results from the fitted linear mixed-effect models on emergence time, plots visited, and swimming time.**

	<i>t</i>	<i>P</i>
Emergence time Females	-0.115	0.909
Emergence time Males	-0.287	0.775
Plots visited Females	-0.384	0.704
Plots visited Males	0.328	0.744
Swimming time Females	-0.558	0.579
Swimming time Males	0.589	0.558