Functional convergence of genomic and transcriptomic genetic architecture underlying sociability in a live-bearing fish

Alberto Corral-Lopez*[1,2,3], Natasha I. Bloch[4], Wouter van der Bijl[1], Maria Cortazar-Chinarro[5,6], Alexander Szorkovszky[7], Alexander Kotrschal[8], Iulia Darolti[1,9], Severine D. Buechel[8], Maksym Romenskyy[10], Niclas Kolm[2], Judith E. Mank[1]

Affiliations

1 - Department of Zoology and Biodiversity Research Centre, University of British Columbia, Vancouver, Canada

2 - Department of Zoology/Ethology, Stockholm University, Stockholm, Sweden.

3 - Division of Ecology and Genetics, Uppsala University, Uppsala, Sweden

4 - Department of Biomedical Engineering, University of Los Andes, Bogota, Colombia

5 - Department of Earth, Ocean and Atmospheric Sciences, University of British Columbia, Vancouver, Canada

6 - MEMEG Department of Biology, Lund University, Lund, Sweden.

7 - RITMO Centre for Interdisciplinary Studies in Rhythm, Time and Motion, University of Oslo, Oslo, Norway

8 - Behavioural Ecology, Wageningen University & Research, Wageningen, Netherlands.

9 - Department of Ecology and Evolution, University of Lausanne, Switzerland

10 - Department of Life Sciences, Imperial College London, London, UK.

*corresponding author: alberto.corral@ebc.uu.se

Abstract

The organization and coordination of fish schools provide a valuable model to investigate the genetic architecture of affiliative behaviors and dissect the mechanisms underlying social behaviors and personalities. We used quantitative genetic methods in replicate guppy selection lines that vary in schooling propensity to investigate the genetic basis of sociability phenotypes. Alignment and attraction, the two major traits forming the sociability personality axis in this species, showed heritability estimates at the upper end of the range previously described for social behaviors, with important variation across sexes. Consistent with findings in collective motion patterns, experimental evolution of schooling propensity increases the sociability of female, but not male, guppies when swimming with unfamiliar conspecifics, a highly relevant trait in a species with fission-fusion social systems in natural populations. Our results from PoolSeq and RNASeq convergently identify genes involved in neuron migration and synaptic function, highlighting a crucial role of glutamatergic synaptic function and calcium-dependent signaling processes in the evolution of schooling.

Introduction

Living in groups, a widespread phenomenon across the animal kingdom, can lead to strikingly complex social behaviors, such as cooperative interactions, subdivision of labor, or collective decision-making[1]. Sociability, the propensity to affiliate with other animals, can also vary across individuals. Sociability represents a fundamental aspect of personality which can influence social interactions and is often subject to strong selective processes[2,3]. Indeed, intra-specific differences in sociability are widespread[e.g 4,5] and often individual genetic variation underlies variability in personality and social behavior phenotypes[6]. However, heritability estimates of social behavior traits are consistent with a complex, polygenic architecture[7]. Human twin and family studies reveal heritability estimates of personality traits generally ≈ 0.40[(reviewed in 8)]. In non-human animals, a meta-analysis estimated the mean heritability is 0.23 across social behaviors, including personality traits[9], with heritability of affiliative associations ranging substantially, from 0.11 to 0.51[10–13].

Despite this complexity, multiple neural and genetic mechanisms underlying social behavior have been identified[14]. Many of the neural structures and neuromodulators (serotonin, dopamine, vasopressin and oxytocin) are highly conserved within the social-decision making network across vertebrates[15]. Moreover, human personality traits associated with social-decision making have been linked to dopaminergic and serotonergic genes[(reviewed in 16)], and the regulation of these neuromodulators has been connected to neurodevelopmental disorders that affect affiliative behaviors, such as autism spectrum disorder[6,17,18]. Studies in non-human organisms likewise point towards a major role of genes involved in the regulation of these neurochemical systems. For instance, mouse knock-out mutants for genes involved in dopaminergic signalling exhibit altered sociability phenotypes[19], and changes in sociability in three-spined sticklebacks, *Gasterosteus aculeatus*, are predicted by natural variation in the expression of genes within the dopaminergic and stress pathways[20]. However, while specific groups of genes have been identified for a range of affiliative behaviors, we lack a deeper understanding of their role in inter-individual variation and evolutionary processes underlying sociability.

In fish, group-living often leads to spectacular forms of collective behavior, with members of a school coordinating their movements to increase efficiency in foraging, travelling or predator

avoidance[1]. The extent to which members of a school coordinate their movements is an integral part of the sociability axis of personality, i.e., how individuals react to the presence or absence of conspecifics excluding aggressive behaviors[21]. We previously showed that schooling behavior has a repeatability of 0.43 at the individual level[22], and that can increase substantially over just three generations of artificial selection in female guppies, *Poecilia reticulata*, generating a 15% increase in intrinsic schooling propensity compared to controls[22,23]. Selection was based on a group phenotype, polarization, or the level of alignment between individuals moving together in a group.

Understanding the genetic basis of this schooling phenotype requires linking individual phenotypic differences to genetic variation. In this study, we phenotyped alignment and attraction of 1496 guppies across 195 families (father, mother, three female and three male offspring from our replicate experimental selection lines) to estimate the heritability of these two motion characteristics that previous factor analyses identified to be integral components for the sociability axis of personality in this species[21]. Because many social interaction patterns in guppies have sex differences[24–26], and because our selection was performed solely on females, we are able to examine cross-sex genetic correlations in this ecologically relevant behavioral trait. Genomic and transcriptomic data from these lines reveal convergence in the genetic architecture of sociability, highlighting a series of genes with well-defined roles in neurodevelopmental processes. Our results provide a robust agreement across experiments about the genetic regulation of neural processes in decision-making and motor control regions of the brain, and its importance for variation of personality within individuals of this species.

Results

Social interactions with unfamiliar individuals and the heritability of sociability in guppies

We first determined whether experimental evolution for higher schooling propensity affected social interactions with unfamiliar conspecifics. For this, we assessed sociability in 740 females and 746 males from multiple families of three replicate lines artificially selected for higher polarization (polarization-selected lines hereon) and three replicate control lines exposed to a group of non-kin unfamiliar conspecifics in an open field test. Specifically, we quantified their alignment and nearest neighbor distance (attraction), two measurements of collective motion

characteristics that are demonstrated to capture the most biologically relevant aspects of the sociability axis of personality in this species[21].

Female guppies from polarization-selected lines presented higher alignment and higher attraction to an unfamiliar group compared to control lines (Linear Mixed Model for alignment, $LMM_{alignment}$: line: $t = 2.27$; $df = 9.68$; $p = 0.047$; $LMM_{attraction}$: line: $t = -2.34$; $df = 9.41$; $p = 0.043$; Fig. 1A; Supplementary Table 1). No differences were observed in these traits between polarization-selected and control males ($LMM_{alignment}$: line: $t = -1.38$; $df = 9.56$; $p = 0.20$; $LMM_{attraction}$: line: $t = 0.88$; $df = 9.26$; $p = 0.40$; Fig. 1A; Supplementary Table 2). Our analyses showed an effect of sex in alignment, with females exhibiting approximately 8% higher alignment than males ($LMM_{alignment}$: sex: $t = -3.02$; $df = 690.08$; $p = 0.003$), but no difference between sexes in attraction to a group of unfamiliar conspecifics ($LMM_{attraction}$: sex: $t = 0.51$; $df = 447.05$; $p = 0.61$; Fig. 1A; Supplementary Table 1). There were some differences in sociability between the parental and offspring generation tested in our experiment, with higher alignment to group average direction and lower distances to nearest neighbor observed in offspring ($LMM_{alignment}$: generation: $t = -10.13$; $df = 1141.24$; $p < 0.001$; $LMM_{attraction}$: generation: $t = 11.29$; $df = 992.16$; $p < 0.001$; Fig. 1A; Supplementary Table 1). Differences in body size between age classes in guppies may explain these results (see Supplementary Table 3), as the time restrictions involved in testing large numbers of fish required that we assessed individuals from the parental and offspring generations at different ages. However, these differences are unlikely to create large biases in our heritability estimates given that we tested all fish after sexual maturation and that polarization-selected and control fish were of similar age within parents tested (9 months old) and within offspring tested (5 months-old). In addition, the difference in means between generations is accounted for in our statistical models (see Methods).

To assess the heritability of sociability in this species, we fitted animal models with alignment and attraction phenotypes quantified from these 1486 individuals, comprised of parents, three male and three female offspring for 195 families (99 polarization-selected and 96 control-families). Given known differences in guppies between the sexes in social interaction patterns[24–26], we estimated heritability with animal models that only included relationships with same sex

167 individuals (same-sex pedigree) or that include relationship with individuals from both sexes (full
168 pedigree).
169
170 Using same-sex pedigree animal models, attraction heritability was similar in females ($h^2_{attraction}$,
171 estimate [95% credible interval] = 0.18 [0.05, 0.34]; Fig. 1B) and males ($h^2_{attraction}$ = 0.19 [0.06,
172 0.34]; Fig. 1B), however alignment heritability was far higher in females ($h^2_{alignment}$ = 0.34 [0.18,
173 0.49]; Fig 1B) than males ($h^2_{alignment}$ = 0.06 [0.00, 0.18]; Fig. 1B). Full-pedigree models indicated
174 lower heritability estimates than those obtained from same-sex pedigree models (Supplementary
175 Table 5; Fig 1B), except for the heritability estimate of attraction in males ($h^2_{attraction}$ = 0.26 [0.16,
176 0.37]; Fig 1B). Finally, animal models indicated a positive female-male genetic correlation in
177 attraction (r_{fm} attraction: 0.68 [0.23, 0.98]; Fig. 1B), although the magnitude of this correlation
178 contained large credible intervals. For alignment, credible intervals for r_{fm} are also wide and span
179 zero (r_{fm} alignment: 0.44 [-0.17, 0.95]), and we can only conclude that the cross-sex genetic
180 correlation is not strongly negative (Supplementary Table 5; Fig 1B).
181
182 *Genetic basis of sociability in guppies*
183
184 Our quantitative genetic analyses of alignment and attraction suggest an important genetic
185 influence in sociability phenotypes of guppies. As such, we sequenced DNA pools (Pool-seq) to
186 identify genome-wide differences in allele frequencies between polarization-selected female
187 guppies that presented high sociability and polarization-selected females that presented low
188 sociability. Specifically, we focused on measurements obtained from females in analyses of
189 alignment to an unfamiliar group. For this, we pooled the DNA from mothers whose family
190 (normalized mother and daughters alignment score; see methods) were in the top 25% and the
191 bottom 25% quartiles from each of the three replicated polarization-selected lines (six total pooled
192 samples with 7 individuals each; Supplementary Figure 1).
193
194 DNA reads were aligned to the guppy reference genome (Guppy_female_1.0 + MT, RefSeq
195 accession: GCA_000633615.2) to compare genome-wide allele frequency differences between
196 high and low sociability guppies. We ran two independent analyses with these aligned sequences.
197 For our first analysis, we merged sequences from the three replicates with high sociability pooled
198 samples and sequences from three replicates with low sociability pooled samples. We filtered

199 merged sequences to 3,004,974 SNPs (see methods) and performed a Fisher's exact test in
200 Popoolation2[27] to identify SNPs that significantly differed in their allele frequencies between
201 guppies with high and low sociability. Using this methodology, we identified 819 SNPs associated
202 with our sociability phenotype (Fisher's exact test, $p < 10^{-8}$; Fig 2A). SNPs over this standard
203 genome-wide significance threshold[28] were mostly found in single physically unlinked positions
204 across the genome, consistent with a polygenic architecture of the trait.
205
206 Out of these 819 significantly different SNPs, 421 were located within genes or gene promoter
207 regions of the guppy genome and were used for further functional characterization in association
208 with Gene Ontology (GO) annotations (273 unique genes). We clustered GO terms based on
209 semantic similarities and found significant overrepresentation of biological process terms related
210 to learning and memory, synaptic functioning, response to stimulus, locomotion and growth (Fig.
211 2B). We likewise found significant overrepresentation of cadherin and calcium-dependent protein
212 binding annotations (molecular components terms; Supplementary Figure 2) and glutamatergic
213 synapse annotations (cellular components terms, Supplementary Figure 3).
214
215 Second, we looked for consistent differences in allele frequencies between high and low sociability
216 pooled samples in our three replicates by performing the Cochran-Mantel-Haenszel test (cmh-test)
217 in Popoolation2[27]. Convergent changes in allele frequency likely represent selected sites, and are
218 less likely the result of genetic drift in any one line. This stringent analysis identified 13 SNPs
219 from ten different chromosomes with consistent significant differences in allele frequencies across
220 the three replicates (cmh-test p-value <0.01 with false discovery rate correction). Five of these
221 SNPs are located within known coding sequence of the guppy genome, of which three are within
222 well-characterized genes in zebrafish and human homologs with important roles for cognitive
223 function – ubiquitin specific peptidase 11 (*usp11*), *supt6* histone chaperone and transcription
224 elongation factor homolog (*supt6h*), and cadherin 13 (*cdh13*; Fig 2A, Table 1). The other two are
225 classified as novel genes, one of them is matched to an RNA-binding protein *Nova-1-like* gene,
226 similarly associated with motor function and changes in synaptic function (Fig 2A, Table 1).
227
228 *Neurogenomic response of schooling in guppies*
229

230 We used transcriptome sequencing to determine differences in gene expression in multiple brain
231 regions of polarization-selected and control females in response to two different social contexts,
232 swimming alone (the "Alone" condition) or schooling in a group (groups of eight unfamiliar
233 females; the "Group" condition). We focused on three separate brain tissues that control distinct
234 functions. The *optic tectum* is involved in sensory processing of visual signals. The *telencephalon*
235 is implicated in decision making. The *midbrain* is associated with motor function in response to
236 auditory and visual stimuli[29,30]. Together, these three brain tissues contain the main components
237 of the social brain network in fish[31,32].
238
239 ***Differential expression analyses.*** We identified genes differentially expressed between lines
240 under each treatment condition and in each brain region separately, to determine the neurogenomic
241 response triggered by schooling in both lines. Gene expression analyses indicated very little
242 overlap in differentially expressed (DE) genes between polarization-selected and control lines
243 (Fig. 3, Supplementary Dataset 1). Specifically, we found that only Adipocyte enhancer-binding
244 protein 2 gene (*AEBP2;* involved in adipocyte differentiation) in the midbrain and an unknown
245 gene in the optic tectum were differentially expressed in both polarization-selected and control
246 lines. Such little overlap suggests that females from different selection lines are activating different
247 transcriptional cascades and biological pathways in response to social context. In polarization-
248 selected lines we found an order of magnitude fewer DE genes in the optic tectum than in the other
249 brain components (n=21 for optic tectum, n=158 for telencephalon, n=109 for midbrain, each p_{adj}
250 < 0.05). Moreover, in the telencephalon and midbrain, DE genes between Alone and Group
251 treatment in the polarization-selected lines were enriched for GO annotations associated with
252 cognition, memory, learning and social behavior (Supplementary Dataset 1). We found enrichment
253 for these annotation terms for DE genes expressed in the optic tectum, but not in the midbrain or
254 telencephalon, of control lines.
255
256 Hierarchical clustering analyses of DE genes showed that females from polarization-selected lines
257 in the Group condition clustered uniquely from polarization-selected females in the Alone
258 condition (Fig. 3, Supplementary Table 6). Similarly, females from polarization-selected lines in
259 the Group condition clustered uniquely from females from control lines under the Group and Alone
260 conditions in both the telencephalon and the midbrain, suggesting a unique response in the regions

261 of the brain associated with behavior to social exposure. This was not observed in samples from
262 the optic tectum, suggesting that visual processing of social treatments did not differ between
263 polarization-selected and control females. Hierarchical clustering analyses using all expressed
264 genes clustered samples by selection line rather than by social context condition (Supplementary
265 Fig. 4), suggesting that social context affects only a targeted subset of the overall transcriptome
266 rather than the majority of genes.
267
268 ***Differential coexpression analyses.*** We used systems biology methods designed to compare the
269 coexpression networks between conditions to identify genes that change in the way they are
270 connected to other genes within the coexpression network across conditions, independent of
271 whether they are differentially expressed[29–31]. Specifically, we used BFDCA[30] to identify
272 differential coexpressed (DC) gene pairs under the Group and Alone conditions (i.e. pairs of genes
273 that significantly change in correlation between the two social contexts for each line[30,32]). Similar
274 to findings in DE analyses, we found little overlap in the genes forming DC gene pairs between
275 comparisons implemented for control and polarization-selected lines (see Supplementary Tables
276 7-8, Fig. 3). Together, our results suggest that polarization-selected lines were activating different
277 biological pathways than control lines to modulate coordinated movement.
278
279 We additionally found a group of genes that are both differentially expressed (DE) and
280 differentially coexpressed (DC) in the same tissue and line, suggesting they might play an
281 important role mediating coordinated movement (Supplementary Tables 8-9). Specifically, in the
282 telencephalon, we identified 4 genes that are both DE and DC in polarization-selected lines:
283 *LRRC24*, *PTPRS, KHDR2, and PP2BA* (Supplementary Table 9). These genes are part of the
284 Calcineurin and the Wnt/Oxytocin signalling pathways, known to be involved in modulating social
285 behavior, learning and memory[33–35]. Enrichment tests confirm the functional relevance of the DC
286 gene pairs identified, revealing an overrepresentation of genes associated with the glutamatergic
287 synapse, as well as visual transduction among DC gene pairs in multiple comparisons
288 (Supplementary Tables 10-11).
289
290 *Functional characterization of genes of interest across experiments*
291

292 We combined the information from our genomic and transcriptomic analyses on polarization-
293 selected and control lines to obtain an intersected delimitation of the gene functions that our
294 analyses highlighted as important in the development and expression of social interactions with
295 conspecifics. Specifically, we used functional analyses in the set of genes with differentiated SNPs
296 between merged sequences of the three replicates with high and low sociability (273 unique genes)
297 as reference, and compared it to functional analyses of genes differentially expressed in three
298 different brain tissues of females following exposure to multiple social conditions. We found a
299 concordance of 79% in the combination of biological processes (BP), cellular components (CC)
300 and molecular functions (MF) GO terms enriched following analyses of differentially expressed
301 genes in the telencephalon (n =158). This value represented a 1.7-fold increase in the concordance
302 of terms in relation to mean values obtained from corresponding enrichment analyses of 1000
303 random sets of 158 genes (see Methods; mean concordance ± [CIs]: 45% [43,47]). We likewise
304 found concordances of 64% for differentially expressed genes in the midbrain (n = 109), and of
305 4.5% for differentially expressed genes in the optic tectum (n = 21). These represented 2.1-fold
306 and 1.1-fold increases in relation to analyses with 1000 random sets of 109 and 21 genes in
307 midbrain and optic tectum respectively (mean concordance midbrain: 30% [28,31]; mean
308 concordance telencephalon: 3.8% [3.6,4.2]). We summarized and visualized GO terms enrichment
309 lists across experiments and tissues sampled using REVIGO[36]. We found a strong overlap between
310 enrichment of GO biological process terms associated with learning and memory, synaptic
311 processes, neuron projection and cell growth, mostly constrained to the telencephalon and
312 midbrain regions (Fig. 4). We found similar patterns in relation to cellular component GO terms,
313 with strong overlap in neuronal components, in particular with high enrichment of terms associated
314 with glutamatergic synapse. Visualization of GO terms associated with molecular functions
315 suggests a major role of genes with protein binding function across experiments, including a role
316 of cadherin-binding related genes in the midbrain (Fig. 4).
317

318 **Discussion**

319
320 We used behavioral phenotyping across guppy families, in conjunction with PoolSeq and RNASeq
321 to identify the genetic architecture of coordinated motion. Our broad range of analyses, spanning
322 genomes, transcriptomes, and phenotypes, provides an exceptional evaluation of the molecular

323 mechanisms underlying sociability in this fish. Our work suggests that genes and gene networks
324 involved in social decision making through neuron migration and synaptic function are key in the
325 evolution of schooling, highlighting a crucial role of glutamatergic synaptic function and calcium-
326 dependent signaling processes.
327
328 Our pedigree-based phenotyping analyses of 195 guppy families from polarization-selected and
329 control lines indicate moderate levels of heritability (Alignment: 0.06 - 0.34; Attraction: 0.09 -
330 0.26), with pronounced sex differences in full pedigree models ($\text{Alignment}_{\text{female-male}}$: 0.10 ± 0.05;
331 $\text{Attraction}_{\text{female-male}}$: -0.17 ± 0.05), in estimates for key behavioral traits forming the sociability axis
332 in this species. Our heritability estimates are similar to estimates for affiliative social behavior
333 traits in primates, ungulates and rodents[10–13,37], and to overall estimates of heritability in
334 personality traits across human and non-human animals[8,38]. Given the importance of social
335 behavior in a range of survival and fitness components in natural systems[1,39,40], our results suggest
336 that complex genetic architectures can respond quickly to strong evolutionary pressures, even
337 when only one sex is subject to selection[22], and that our lab population contained significant
338 amounts of standing genetic variation for these traits prior to selection.
339
340 The complex genetic architecture makes it difficult to precisely characterize cross sex genetic
341 effects in our study. We nonetheless observed a positive cross sex genetic correlation in attraction
342 (0.68, CI: 0.25 - 0.98), suggesting similarities between males and females in the genetic
343 architecture of this trait. This result is concordant with a study focused on the bold-shy continuum
344 aspect of personality establishing that sex differences in risk-taking behaviors are weak and likely
345 lack sex-specific genetic architecture in this species[41]. Yet, sex differences in heritability estimates
346 of alignment (♀$h^2_{\text{alignment}}$ = 0.34 [0.18, 0.49]; ♂$h^2_{\text{alignment}}$ = 0.06 [0.00, 0.18]) suggest that it is
347 important to account for sex-specific additive genetic variance when inferring the evolvability of
348 personality traits. Additionally, the low cross-sex heritability we observe in these latter traits is
349 particularly interesting, and suggests that selection in one sex for a complex trait need not result in
350 a correlated response in the other sex. Overall, this indicates significant sex-specific genetic
351 variation for sex-specific behaviours, and that sexually dimorphic behaviors need not require
352 decoupling of male and female genetic architecture when sufficient sex-specific genetic variation
353 is present.

354

355 We next mapped the genomic and transcriptomic basis of phenotypic differences in polarization
356 in female guppies. Our GWAS experimental design was designed to compare individuals with
357 high- and low-sociability phenotypes from within polarization-selected lines, rather than between
358 polarization-selected and control lines. This may have resulted in compressed phenotypic spread,
359 but carries the important advantage of reducing the incidence of SNP frequencies that vary across
360 alternative selection lines due to drift. As such, our design is conservative. In our most stringent
361 PoolSeq analysis, we identified SNPs in four genes that consistently differed across all three
362 replicate lines, which have previously been associated with cognition and functions relevant to
363 social behavior. The supt6 histone chaperone and transcription elongation factor homolog (*supt6h*)
364 is important in the positive regulation of transcriptional elongation and a substrate of mTOR, a
365 signalling pathway with a role in cognitive function[42,43]. Ubiquitin specific peptidase 11 (*usp11*)
366 homolog in humans has a critical function in the development of neural cortex, and knock-out
367 studies in mice show that the locus protects females from cognitive impairment[44,45]. Similarly, the
368 novel RNA-binding protein *Nova-1-like* gene is associated with a neuron-specific nuclear RNA
369 binding protein in humans and regulates brain-specific splicing related to synaptic function[46].
370

371 Finally, our PoolSeq analysis identified Cadherin 13 (*cdh13*), the human homolog of which has a
372 crucial role in GABAergic function[47], with involvement in neural growth and axonal guidance
373 during early development[48,49]. Moreover, deficit of this gene has a major impact in
374 neurodevelopmental disorders including attention-deficit/hyperactivity disorder and autism
375 spectrum disorder[50]. Indeed, *cdh13* knock-out mice display delayed acquisition of learning tasks
376 and a decreased latency in sociability[51]. Interestingly, repeated selection of genes involved in
377 cadherin-signalling pathways[52] has been shown in guppy populations experiencing different
378 predation pressures. Together, natural selection imposed by differences in predation across these
379 populations[53–55], and our combined findings in the genomic background of guppies suggest strong
380 selective pressures for cadherin-signalling genes due to their modulation of affiliative behaviors.
381

382 Our expression results revealed differences in regulation in genes associated with learning,
383 behavior, and neural function, mainly in the telencephalon and midbrain, in comparisons of
384 polarization-selected and control lines in different social contexts. Overall, this suggests that the

385 regulation of highly demanding cognitive processes via synaptic function underlies variation in
386 sociability. While the integration of visual signals is central in fish schools[56], our results suggest
387 that higher order cognitive processes are the basis of variation in social affinity. Indeed, the
388 differences in alignment and attraction observed when swimming with unfamiliar conspecifics are
389 arguably highly cognitively demanding, as within a collective motion context, the tendency to
390 copy the directional movements of other individuals implies a direct trade-off between personal
391 goal-oriented behaviors and the benefits of social conformity[57,58]. Together, our study of
392 transcriptomic profiles of schooling fish suggests that the regulation of affiliative behaviors in this
393 species is driven by an intricately linked social decision-making network in the brain[59], with strong
394 links to functional groups governing social behaviors and personality across species. More
395 broadly, our results offer insight into important questions about the evolution of behaviour, and
396 other traits with complex genetic architecture. First, our results of large-scale expression
397 differences among selection lines are consistent with recent discussions of the role of gene
398 regulatory networks in co-ordinating large numbers of genes associated with behaviours[60]. It is
399 highly likely that the genes with convergent expression changes in the selection lines are controlled
400 via a modular regulatory architecture, as evidenced by our coexpression network analysis
401 (Supplementary Tables 7-8, Fig. 3).
402
403 We find a striking concordance in the functionality of genes independently identified in genomic
404 and transcriptomic profiling of strongly differentiated experiments assessing social interactions of
405 polarization-selected female guppies (Fig. 4). The overlap in significantly enriched Gene Ontology
406 terms, including learning, synaptic processes and neuron projection restricted to brain regions
407 associated with decision-making and motor control, strongly reinforces the notion that genetic
408 regulation of these cognitive processes is fundamental for sociability in guppies. Additionally, the
409 functional concordance we observe between the regulatory and protein differences among our
410 selection lines is noteworthy in the context of the discussion of whether structural or regulatory
411 variation is more important in adaptive phenotypes[61,62]. The overlap in functionality in our
412 genomic and transcriptomic approaches suggests that both are important, with artificial selection
413 for behavior acting on coding and regulatory variation within the same pathway to achieve adaptive
414 phenotypes.
415

416 Our results indicate that the regulation of glutamatergic synaptic processes is a particularly
417 promising network for future studies of affiliative behaviors. Interestingly, differential gene
418 expression of glutamate receptor genes has been identified to regulate female mating preferences
419 in guppies[63], and is concordantly identified across species of vertebrates in the regulation of long-
420 term affiliative mating behaviors[64]. Guppies are livebearers, and this has hindered the use of
421 functional genetic tools such as CRISPR on this species. Although not feasible at this time, future
422 functional validation via genetic manipulations of guppies of these pathways would prove
423 extremely interesting.
424
425 Our results are also concordant with other comparative transcriptomic studies of behavioral
426 responses towards conspecific territorial intrusions, which identified calcium ion-binding
427 regulation across phylogenetically distant species[65]. Together, the consistency in our findings of
428 specific genes and functional terms associated with calcium-dependent and cadherin binding
429 molecular functions across our experiments suggest these are promising molecular targets for
430 future research exploring the evolution and regulation of sociability and affiliative behaviors.
431
432 **Methods**
433
434 Study system
435 To evaluate the genetic architecture of sociability, we performed a series of experiments in guppies
436 following artificial selection on coordinated motion. The laboratory population of guppies used
437 originated from a downstream population of the Quare river in Trinidad, which is subject to high
438 predation levels. The original collection was made in 1998[66], and the laboratory population has
439 since been kept in several large (>200-litre) tanks of >200 individuals each to avoid inbreeding.
440 The artificial selection procedure is outlined in detail in[22,23]. In brief, groups of female guppies
441 were subjected to repeated open field tests, and were subsequently sorted based on their median
442 polarization, measured by the degree of alignment exhibited by the individuals within the group
443 when swimming together[22,23]. For three generations, females from groups with higher polarization
444 were mated with males from those cohorts to generate three lines of guppies that had been selected
445 for high polarization. In parallel, random females were exposed to the same experimental
446 conditions and were mated with unselected males to generate three control lines. Analysis of the
447 third generation of polarization-selection revealed that, on average, females exhibited a 15%

448 higher level of polarization and a 10% higher level of group cohesiveness compared to control
449 females[22].
450
451 Throughout the selection experiment and the completion of experiments described below all fish
452 were removed from their parental tanks after birth, separated by sex at the first onset of sexual
453 maturation, and afterwards kept in single-sex groups of eight individuals in 7-L tanks containing
454 2 cm of gravel with continuously aerated water, a biological filter, and plants for environmental
455 enrichment. We allowed for visual contact between the tanks. The laboratory was maintained at
456 26°C with a 12-h light:12-h dark schedule. Fish were fed a diet of flake food and freshly hatched
457 brine shrimp daily.
458
459 Social interactions with unfamiliar individuals and the heritability of sociability
460 To investigate the heritability and cross-sex genetic correlations of sociability in the guppy, we
461 measured alignment and attraction with unfamiliar groups of conspecifics in parents and offspring
462 from polarization-selected and control lines. Specifically, using offspring of the F3 generation of
463 selection we bred 35 families for each of the three polarization-selected and for each of the three
464 control lines. From our population of F3 generation offspring (kept in single sex groups prior to
465 the breeding experiments), we used male and female guppies of the same age (approximately 9
466 months old) and paired them in 3L tanks to generate the parental generation. We collected
467 offspring for the two first clutches of these pairs and we transferred newborn offspring to 3L tanks
468 in groups of three siblings. We separated siblings by sex at the first onset of sexual maturation,
469 and afterwards we kept them in single-sex groups of three individuals until behavioral testing. We
470 phenotyped sociability of a total of 195 guppy families: mother, father and six offspring (three
471 females and three males). Any family for which we did not collect at least three female and three
472 male offspring was disregarded from further behavioral testing. Each of the six selection lines was
473 represented with a minimum of 30 families in our heritability analyses.
474
475 ***Behavioral assays.*** To phenotype sociability in each member of our guppy families, we measured
476 alignment and attraction of 1495 guppies from our breeding experiment. For each fish, we
477 performed an open field assay using white arenas with 55 cm diameter and 3 cm water depth in
478 which our focal fish (guppies from the breeding experiment) interacted with a group of seven

479 same-sex conspecifics. Non-focal guppies used in these assays were from a lab wild-type stock
480 population and of similar age to our focal fish. Prior to the start of the test, focal fish and the seven-
481 fish group were acclimated in the centre of the arena for one minute in separate opaque white 15
482 cm PVC cylinders. After this acclimation period, we lifted the cylinders and filmed the arena for
483 10 minutes using a Point Grey Grasshopper 3 camera (FLIR Systems; resolution, 2048 pixels by
484 2048 pixels; frame rate, 25 Hz). Three weeks prior to assays, we tagged wild-type fish with small
485 black elastomere implants (Northwest Marine Technology Inc.) to allow recognition of wild-type
486 fish after completion of each assay. After completion, we gently euthanized focal fish from the
487 parental generation with an overdose of benzocaine and kept them in ethanol for future genomic
488 analyses. Focal fish from the offspring generation were transferred to group tanks for future
489 experimental use. Groups of seven wild-type fish were transferred to holding tanks and used in a
490 maximum of seven assays with focal fish.
491
492 ***Data processing.*** We tracked the movement of fish groups in the collected video recordings using
493 idTracker[67] and used fine-grained tracking data to calculate the following variables in Matlab
494 (v2020): i) alignment, the median alignment of the focal fish to the group average direction across
495 all frames in the assay. This was quantified by the total length of the sum of two-unit vectors, one
496 representing the heading of the focal fish, and one representing the heading of the group centroid.
497 Calculations of alignment were only obtained if six out of the eight members of the group presented
498 tracks following the optimization of our tracking protocol in the setup in ([22,23,68]); ii) attraction, the
499 median nearest neighbor distance across all frames in the assay; and iii) activity, we obtained the
500 median speed across all group members and across all frames by calculating the first derivatives
501 of the x and y time series, then smoothed using a third-order Savitzky–Golay filter. For all
502 measurements, trials with less than 70% complete tracks (n= 8) were disregarded for further
503 analyses. The proportion of frames used did not differ between polarization-selected and control
504 fish for any comparison across different generations and sexes (Supplementary Figure 5). We
505 calculated these variables for the focal fish and the average for the seven-fish wild-type group. To
506 recover focal fish id in the tracking data we used idPlayer to visualize trials by projecting the raw
507 tracking data onto experimental videos. We followed focal individuals for the first two minutes of
508 the assay and used the stable identity assigned by idTracker in data collection. In trials with less
509 than 85% complete tracks (n= 8), we followed focal individuals for the total duration of the

510 recording to verify the consistency of identity assigned by idTracker. This approach has previously
511 shown strong reliability in individuals that were observed using this protocol for 20-min recordings
512 in the same experimental setup that quantified sexual behavior of guppies in mixed-sex shoals[69].
513
514 ***Statistical analyses.*** Analyses were conducted using R Statistical Software (v4.1.3)[70], RStudio
515 (v2023.3.1.446)[71] and the tidyverse package[72]. We used Linear Mixed Models (LMMs) with
516 alignment and attraction as dependent variables to test for potential differences between
517 polarization-selected and control lines in social interactions with unfamiliar individuals. Selection
518 regime, sex, the interaction between these two factors and generation were included as fixed
519 effects. The average activity of the wild-type group was coded as a covariate, with a random
520 intercept for each replicated selection line, the breeding family, and the number of tests previously
521 performed with the wild-type group as random factors. All models were run using lme4 and
522 lmerTest packages[73,74]. Model diagnostics showed that residual distributions were roughly normal
523 with no evidence of heteroscedasticity.
524
525 To estimate heritability, the degree of phenotypic variation due to genetic inheritance, and cross-
526 sex genetic correlations of alignment and attraction we used Bayesian animal models[75]. Animal
527 models use a matrix of pedigree relationships set as a random effect, to separate phenotypic
528 variance for each response variable into additive genetic variance and the remaining variance.
529 Given strong sex differences in social interactions in guppies, we performed three animal models
530 for each trait: one including the data on the 1495 phenotyped individuals, and two including only
531 the phenotyped females or males. Parameter values were estimated using the brms interface[76,77] to
532 the probabilistic programming language Stan[78]. We used normal priors with a mean of 0 and
533 standard deviation (sd) of 3 for fixed effects, and student-t priors with 5 degrees of freedom, a
534 mean of 0 and sd of 5 for random effects. The full pedigree model estimated cross-sex correlations
535 with a Lewandowski-Kurowicka-Joe (LKJ) prior with $\eta = 1$, which is uniform over the range -1
536 to 1. Posterior distributions for full / same-sex pedigree models were obtained using Stan's no-U-
537 turn HMC with 24 / 16 independent Markov chains of 2500 / 4000 iterations, discarding the first
538 1500 / 2000 iterations per chain as warm-up and resulting in 24000 / 32000 posterior samples
539 overall. Convergence of the chains and sufficient sampling of posterior distributions were
540 confirmed by a potential-scale-reduction metric (R) below 1.01, and an effective sample size of at

541 least 1000. For each model, posterior samples were summarized based on the Bayesian point
542 estimate (posterior median) and posterior uncertainty intervals by HDIs. We calculated estimates
543 of heritability by taking the ratio of the additive genetic variance to the total phenotypic variance
544 in each independent model (see Supplementary Tables 5-6).
545
546 Genetic basis of sociability in guppies
547
548 ***Pooled DNA sequencing.*** We extracted DNA of muscle tissue from the caudal peduncle of
549 polarization-selected females from the parental generation using Qiagen's DNAeasy Blood &
550 Tissue kit following standard manufacturer's protocol with an additional on-column RNase A
551 treatment. We quantified DNA concentration using fluorometry (Qubit; ThermoFisher Scientific).
552 We next pooled samples from the seven females that represented the top and bottom 20%
553 polarization-selected guppy lines whose families presented higher and lower sociability in six final
554 pools at equimolar amounts (Supplementary Figure 1). We achieved a minimum of three µg
555 genomic DNA per pool. We used a Nextera DNA Flex library preparation kit (Illumina) following
556 the manufacturer protocol. The final library containing six pooled samples was sequenced at
557 SciLife Lab, Uppsala (Sweden) in one lane of an Illumina NovaSeq 6000. We obtained on average
558 31.8 million 150bp read pairs per sample (26.9 million read pairs minimum per sample).
559
560 ***Read quality control and trimming.*** We assessed the quality of reads for each pool using FastQC
561 v.0.11.4 (www.bioinformatics.babraham.ac.uk/projects/fastqc). After verifying initial read
562 quality, reads were trimmed with Trimmomatic v0.35[79]. We filtered adaptor sequences and
563 trimmed reads if the sliding window average Phred score over four bases was <15 or if the
564 leading/trailing bases had a Phred score <4, removing reads post filtering if either read pair was
565 <50 bases in length. Quality was verified after trimming with FastQC.
566
567 ***Genome-wide allele frequency analysis.*** Reads were mapped using default settings to the guppy
568 reference genome assembly (Guppy_female_1.0 + MT, RefSeq accession: GCA_000633615.2)[80]
569 with bwa-mem (v0.7.17)[81]. We used Samtools (v. 1.6.0)[82] to convert sam to bam files, to sort bam
570 files, to remove duplicates, and to make mpileup files. First, to identify SNPs that significantly
571 differed in their allele frequencies between guppies with high and low sociability, we merged
572 sequences from high sociability and low sociability pools and used Popoolation2[27] to create a

573 synchronized file with allele frequencies for high and low sociability (mpileup2sync.pl –min-qual
574 20), compute allele frequency differences (mpileup2sync.pl --min-count 6 --min-coverage 25 --
575 max-coverage 200), calculate Fst for every SNP (fst-sliding.pl) and perform a Fisher's exact test
576 (fisher-test.pl). Second, we similarly used Popoolation2 to detect consistent changes in allele
577 frequencies of sociability pooled samples for our three replicated artificial selection lines. For this,
578 we created one sync file per replicate (mpileup2sync.pl –min-qual 20), and perform a Cochran-
579 Mantel-Haenszel (CMH) test (cmh-test.pl --min-count 18 --min-coverage 25 --max-coverage
580 200). Using package qqman[83] in R (v 4.1.3)[70] we made Manhattan plots for each chromosome by
581 plotting the negative log10-transformed p-values of exact fisher and CMH tests as a function of
582 chromosome position.
583
584 ***Significance tests and functional analyses.*** We determined SNPs that were significantly different
585 between high and low sociability merged pools in Fisher's exact tests using the traditional genome-
586 wide significance threshold [$-\log10(p) > 8$][28]. We next used custom scripts to identify the overlap
587 between position of these SNPs and genes present in the guppy reference annotated genome[80], and
588 to find homologous genes of these set in medaka (*Oryzias latipes*). We further used this set of
589 unique genes (n = 160) to determine associated GO terms between our merged pools. For this, we
590 performed enrichment tests in PANTHER[84], as implemented in the GO Ontology Consortium
591 (http://www.geneontology.org/). To test for enrichments of GO terms we performed one-tail
592 Fisher's exact tests with a Bonferroni corrected p-value threshold of $p < 0.05$ using a full list of
593 medaka genes orthologous to guppy genes as background. We used Revigo (http://revigo.irb.hr)[36]
594 to find and visualize representative subsets of terms based on semantic similarity measurements
595 for our enriched GO terms related to biological processes, cellular components and molecular
596 functions.
597
598 For CMH test results, we determined SNPs that were significantly different between high and low
599 sociability pools based on FDR corrected p-value < 0.01. We used a custom script to identify the
600 overlap between position of these SNPs and genes present in the guppy reference annotated
601 genome[80].
602
603 <u>Neurogenomic response of schooling in guppies</u>
604

605 ***Behavioral assays and tissue collection.*** Using offspring of the F3 generation (6-months-old), we
606 placed an individual or groups of eight unfamiliar adult control and polarization-selected females
607 in white 55cm arenas. After 30 min females were euthanized by transfer to ice water. After 30
608 seconds, with aid of a Leica S4E microscope, we removed the top of the skull and after cutting
609 transversally posterior of the optic tectum and anterior of the cerebellum, and horizontally through
610 the optic chiasm, removed the brain from the skull and placed it into ice water. We severed the
611 *telencephalon* from the rest of the brain between the ventral telencephalon and thalamus at the
612 'commissura anterioris', including both the pallium and subpallium regions. Then we cut the
613 laminated cup-like structures of the *optic tectum*. The remaining part of the brain was the *midbrain*.
614 Dissections took under 2 minutes and tissue samples were immediately preserved in RNAlater
615 (Ambion) at 4°C for 24 hours and then at -20°C until RNA extraction.
616
617 ***RNA extraction and sequencing.*** For each treatment, we pooled tissue from ten individuals into
618 two non-overlapping pools of five for each replicate line. We used this strategy to reduce noise in
619 transcript expression data during sample normalization procedures potentially caused by outliers
620 during behavioral experiments, while maintaining each replicate as a comparable unit. Our
621 experimental design represents a total of 120 individual females, constituting six pools per
622 treatment, per selection regime for a total of 24 pools per tissue. Each sample pool was
623 homogenized, and RNA was extracted using Qiagen's RNAeasy kits following standard
624 manufacturer's protocol. Libraries for each sample were prepared and sequenced by the Wellcome
625 Trust Center for Human Genetics at the University of Oxford, UK. All samples were sequenced
626 across nine lanes on an Illumina HiSeq 4000. We obtained on average 33.9 million 75bp read pairs
627 per sample (28.9 million read pairs minimum, 39.8 million maximum).
628
629 ***Read quality control and trimming.*** We assessed the quality of reads for each sample using
630 FastQC v.0.11.4 (www.bioinformatics.babraham.ac.uk/projects/fastqc). After verifying initial
631 read quality, reads were trimmed with Trimmomatic v0.35[79]. We filtered adaptor sequences and
632 trimmed reads if the sliding window average Phred score over four bases was < 15 or if the
633 leading/trailing bases had a Phred score < 3, removing reads post filtering if either read pair was <
634 33 bases in length. Quality was verified after trimming with FastQC.
635

636 ***Differential Expression Analysis.*** We mapped RNAseq reads against the latest release of the
637 published guppy genome assembly[80], using the HiSat 2.0.5 – Stringtie v1.3.2 suite[81]. For each
638 individual pool, reads were mapped to the genome and built into transcripts using default
639 parameters. The resulting individual assemblies were then merged into a single, non-redundant
640 assembly using the built-in StringTie-merge function. We filtered the resulting assembly for non-
641 coding RNA using medaka and Amazon molly (*Poecilia formosa*) non-coding RNA sequences as
642 reference in a nucleotide BLAST (Blastn). After eliminating all sequences matching non-coding
643 RNAs, we kept only the longest isoform representative for each transcript for further analysis.
644 Finally, we quantified expression by re-mapping reads to this filtered assembly using RSEM
645 (v1.2.20)[85].
646
647 Lowly expressed genes were removed by filtering transcripts < 2 RPKM (reads per kilobase per
648 million mapped reads), preserving only those transcripts that have expression above this threshold
649 in at least half of the samples for each treatment within a line. After this final filter, a total of
650 26,140 optic tectum transcripts, 25,100 telencephalon transcripts and 26,514 midbrain transcripts
651 were retained for further analysis. Using sample correlations in combination with MDS plots based
652 on all expressed transcripts, we determined that none of the 72 pools represented outliers and thus,
653 all samples were included in the analysis.
654
655 We used DESeq2[86] to normalize filtered read counts using standard function to identify
656 differentially expressed (DE) genes between the alone and the group treatment in control and
657 polarization-selected lines separately, and then examined the overlap in differentially expressed
658 genes between them. A transcript was considered differentially expressed if it had an FDR
659 corrected p-value < 0.05. As behavior could be modulated by small changes in expression, we did
660 not filter differentially expressed genes based on Log Fold-change in expression between the
661 treatments.
662
663 ***Differential Coexpression Analysis.*** We used Bayes approach for Differential Coexpression
664 Analysis (BFDCA)[30], in order to identify pairs of genes that have different correlation patterns in
665 two conditions[32,87,88]. Here we compared the Alone and Group treatments within each line for each
666 tissue separately, in the same manner as the previously described DE analysis. BFDCA is based

667 on WGCNA and has been shown to be a reliable and accurate method[30]. This untargeted approach
668 to differential coexpression (DC) analysis uses a combined Bayes factor, a ratio of marginal
669 likelihood of the data between the two alternative hypotheses, to evaluate which genes are
670 differentially correlated in two conditions. We controlled for false positives and accounted for
671 multiple testing by integrating a random permutations approach[32]. In short, we created 1000
672 permuted datasets and considered a DC gene pair significant if the Bayes factor for the actual
673 expression data was larger than the 1% tail of the permutated data Bayes factor distribution.
674
675 ***Functional Analyses***. To investigate the function of DE genes we performed GO term enrichment
676 tests. To accomplish this, we initially completed the annotation of the reference genome assembly.
677 The transcripts without clear gene names from the reference genome, and the de novo transcripts
678 identified by HiSat were annotated with blastX against the Swissprot non-redundant database. We
679 then determined which GO terms were associated with differentially expressed genes and
680 performed Biological Processes (BP), Cellular Components (CC) and Molecular Functions (MF)
681 enrichment tests in PANTHER[84]. To assess the level of concordance between genes of interest
682 across experiments, we compared the proportion of BP, CC and MF GO terms that were
683 significantly enriched in genomic analyses of sociability implemented in polarization-selected
684 females, and the proportion of BP, CC and MF GO terms enriched in differential expression
685 analyses in brain tissue of polarization-selected females following exposure to Group and Alone
686 experimental conditions. To assess their significance, we compared these values to mean
687 proportions obtained from bootstrap analyses of 1000 random sets of 158 (for comparison with
688 telencephalon), 109 (midbrain) and 21 (optic tectum) genes from our medaka-guppy orthologous
689 gene list. All analyses were based on one-tail Fisher's exact tests with a Bonferroni corrected p-
690 value threshold of $p < 0.05$ using medaka genes orthologous to guppy genes as the background.
691 Bootstrap analyses with random sets of genes were automated using rbioapi package[89] in R
692 (v4.1.3)[70]. We next summarized and visualized GO terms enrichment lists across experiments and
693 tissues using REVIGO[36] (settings: SimRel semantic similarity measure, 0.5 value). To investigate
694 the function of differentially coexpressed genes, we used g:Profiler[90] to identify the enriched BP
695 GO terms and pathways that were altered across mating contexts associated with differentially
696 coexpressed gene pairs. We determined over-represented pathways among DC gene pairs in each

tissue using the human (Homo sapiens) database in g:Profiler. We chose the human database for its completeness, acknowledging the distant phylogenetic relationship to guppies.

Acknowledgements

We thank David Sumpter, Kristiaan Pelckmans and James Herbert-Read for important contributions to the conceptualization of the artificial selection procedure. We thank three anonymous reviewers for comments and suggestions to revise a previous version of the manuscript. We are grateful to Jacelyn Shu for the design of guppy graphics for figures. We thank Anna Rennie, Emilio Trejo, and Annika Boussard for help with fish husbandry. **Funding**: This work was supported by the Knut and Alice Wallenberg Foundation (102 2013.0072 to N.K.), the Canada 150 Research Chair Program, and the European Research Council (680951 to J.E.M), the Swedish Research Council (2016-03435 to N.K., 2017-04957), The Royal Swedish Academy of Sciences (BS2019-0046 to A.C-L.), Lars Hiertas Memorial Foundation (FO2019-0477 to A.C-L.), European Research Council; H2020 Marie Skłodowska-Curie Actions, (654699 to N.I.B); Universidad de los Andes (FAPA-4700000443 to N.I.B). **Author contributions**: J.E.M., N.K. and A.C-L. contributed to conceptualization and funding acquisition of the project. N.K., A.K., A.S., and M.R. contributed to the design of the selection procedure and behavioral experiments. A.C-L. conducted research to obtain behavioral data. A.C-L. and W.vdB. performed formal analyses and visualization of behavioral and heritability data. A.C-L., M.C-C. conducted research to obtain genomic data. A.C-L., M.C-C. and I.D. perform formal analyses and visualization of genomic data. N.B. and A.K. conducted research to obtain transcriptomics data. N.B. and. A.C-L. performed formal analyses and visualization of transcriptomics data. A.C-L., J.E.M., and N.K. wrote the original draft. All authors contributed to the final version of the manuscript. **Competing interests**: The authors declare that they have no competing interests. **Data and materials availability**: Data and code needed to support the conclusions in the paper are deposited in figshare repository (doi: 10.6084/m9.figshare.23805702). Genomic and transcriptomic data is deposited at NCBI (accession codes PRJNA994132 and PRJNA504011). Additional data related to this paper may be requested from the authors.

References

1. Krause, J. & Ruxton, G. D. *Living in Groups*. (Oxford University press, 2002).
2. Réale, D., Reader, S. M., Sol, D., McDougall, P. T. & Dingemanse, N. J. Integrating animal temperament within ecology and evolution. *Biological Reviews* **82**, 291–318 (2007).

3. Gartland, L. A., Firth, J. A., Laskowski, K. L., Jeanson, R. & Ioannou, C. C. Sociability as a personality trait in animals: methods, causes and consequences. *Biological Reviews* **97**, 802–816 (2022).
4. Aplin, L. M. *et al.* Individual personalities predict social behaviour in wild networks of great tits (Parus major). *Ecol Lett* **16**, 1365–1372 (2013).
5. Bevan, P. A., Gosetto, I., Jenkins, E. R., Barnes, I. & Ioannou, C. C. Regulation between personality traits: individual social tendencies modulate whether boldness and leadership are correlated. *Proceedings of the Royal Society B* **285**, (2018).
6. Ebstein, R. P., Israel, S., Chew, S. H., Zhong, S. & Knafo, A. Genetics of Human Social Behavior. *Neuron* **65**, 831–844 (2010).
7. Robinson, G. E., Fernald, R. D. & Clayton, D. F. Genes and social behavior. *Science (1979)* **322**, 896–900 (2008).
8. Turkheimer, E., Pettersson, E. & Horn, E. E. A Phenotypic Null Hypothesis for the Genetics of Personality. *Annual Reviews in Psychology* **65**, 515–540 (2014).
9. Dochtermann, N. A. *et al.* The Heritability of Behavior: A Meta-analysis. *Journal of Heredity* 1–8 (2019).
10. Brent, L. J. N. *et al.* Genetic origins of social networks in rhesus macaques. *Scientific Reports 2013 3:1* **3**, 1–8 (2013).
11. Lea, A. J., Blumstein, D. T., Wey, T. W. & Martin, J. G. A. Heritable victimization and the benefits of agonistic relationships. *Proc Natl Acad Sci U S A* **107**, 21587–21592 (2010).
12. Staes, N. *et al.* Bonobo personality traits are heritable and associated with vasopressin receptor gene 1a variation. *Scientific Reports 2016 6:1* **6**, 1–8 (2016).
13. Knoll, A. T., Jiang, K. & Levitt, P. Quantitative trait locus mapping and analysis of heritable variation in affiliative social behavior and co-occurring traits. *Genes Brain Behav* **17**, (2018).
14. Fisher, D. N. Direct and indirect phenotypic effects on sociability indicate potential to evolve. *J Evol Biol* **36**, 209–220 (2023).
15. Fisher, D. N. & McAdam, A. G. Social traits, social networks and evolutionary biology. *J Evol Biol* **30**, 2088–2103 (2017).
16. O'Connell, L. A. & Hofmann, H. A. Genes, hormones, and circuits: An integrative approach to study the evolution of social behavior. *Front Neuroendocrinol* **32**, 320–335 (2011).
17. Balestri, M., Calati, R., Serretti, A. & De Ronchi, D. Genetic modulation of personality traits: A systematic review of the literature. *Int Clin Psychopharmacol* **29**, 1–15 (2014).
18. Sahin, M. & Sur, M. Genes, circuits, and precision therapies for autism and related neurodevelopmental disorders. *Science (1979)* **350**, (2015).
19. Shou, J., Tran, A., Snyder, N., Bleem, E. & Kim, S. Distinct Roles of GluA2-lacking AMPA Receptor Expression in Dopamine D1 or D2 Receptor Neurons in Animal Behavior. *Neuroscience* **398**, 102–112 (2019).
20. Abbey-Lee, R. N., Kreshchenko, A., Fernandez Sala, X., Petkova, I. & Løvlie, H. Effects of monoamine manipulations on the personality and gene expression of three-spined sticklebacks. (2019).
21. Sumpter, D. J. T., Szorkovszky, A., Kotrschal, A., Kolm, N. & Herbert-Read, J. E. Using activity and sociability to characterize collective motion. *Philosophical Transactions of the Royal Society B: Biological Sciences* **373**, 20170015 (2018).

22. Kotrschal, A. *et al.* Rapid evolution of coordinated and collective movement in response to artificial selection. *Sci Adv* **6**, eaba3148 (2020).
23. Szorkovszky, A. *et al.* An efficient method for sorting and quantifying individual social traits based on group-level behaviour. *Methods Ecol Evol* **8**, 1735–1744 (2017).
24. Piyapong, C. *et al.* Sex matters: a social context to boldness in guppies (Poecilia reticulata). *Behavioral Ecology* **21**, 3–8 (2010).
25. Dimitriadou, S., Croft, D. P. & Darden, S. K. Divergence in social traits in trinidadian guppies selectively bred for high and low leadership in a cooperative context. doi:10.1038/s41598-019-53748-4.
26. Griffiths, S. W. & Magurran, A. E. Sex and schooling behaviour in the Trinidadian guppy. *Anim Behav* **56**, 689–693 (1998).
27. Kofler, R., Pandey, R. V. & Schlötterer, C. PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics* **27**, 3435–3436 (2011).
28. Pe'er, I., Yelensky, R., Altshuler, D. & Daly, M. J. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genet Epidemiol* **32**, 381–385 (2008).
29. Iancu, O. D., Colville, A., Darakjian, P. & Hitzemann, R. Coexpression and Cosplicing Network Approaches for the Study of Mammalian Brain Transcriptomes. *Int Rev Neurobiol* **116**, 73–93 (2014).
30. Wang, D., Wang, J., Jiang, Y., Liang, Y. & Xu, D. BFDCA: A Comprehensive Tool of Using Bayes Factor for Differential Co-Expression Analysis. *J Mol Biol* **429**, 446–453 (2017).
31. Jiang, Z., Dong, X., Li, Z.-G., He, F. & Zhang, Z. Differential Coexpression Analysis Reveals Extensive Rewiring of Arabidopsis Gene Coexpression in Response to Pseudomonas syringae Infection OPEN. *Nature Publishing Group* (2016) doi:10.1038/srep35064.
32. Bloch, N. I. *et al.* Different mating contexts lead to extensive rewiring of female brain coexpression networks in the guppy. *Genes Brain Behav* **20**, e12697 (2021).
33. Froemke, R. C. & Young, L. J. Oxytocin, Neural Plasticity, and Social Behavior. *https://doi.org/10.1146/annurev-neuro-102320-102847* **44**, 359–381 (2021).
34. Miyakawa, T. *et al.* Conditional calcineurin knockout mice exhibit multiple abnormal behaviors related to schizophrenia. *Proc Natl Acad Sci U S A* **100**, 8987–8992 (2003).
35. Lee, J. *et al.* Opposing Functions of Calcineurin and CaMKII Regulate G-protein Signaling in Egg-laying Behavior of C. elegans. *J Mol Biol* **344**, 585–595 (2004).
36. Supek, F., Bošnjak, M., Škunca, N. & Šmuc, T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* **6**, (2011).
37. Albery, G. F. *et al.* Multiple spatial behaviours govern social network positions in a wild ungulate. *Ecol Lett* **24**, 676–686 (2021).
38. Dochtermann, N. A., Schwab, T. & Sih, A. The contribution of additive genetic variation to personality variation: heritability of personality. *Proceedings of the Royal Society B: Biological Sciences* **282**, (2015).
39. Couzin, I. D., Krause, J., Franks, N. R. & Levin, S. A. Effective leadership and decision-making in animal groups on the move. *Nature* **433**, 513–516 (2005).
40. Parrish, J. K. & Edelstein-Keshet, L. Complexity, Pattern, and Evolutionary Trade-Offs in Animal Aggregation. *Science (1979)* **284**, 99–101 (1999).

41. White, S. J., Houslay, T. M. & Wilson, A. J. Evolutionary genetics of personality in the Trinidadian guppy II: sexual dimorphism and genotype-by-sex interactions. *Heredity 2018 122:1* **122**, 15–28 (2018).
42. Lipton, J. O. & Sahin, M. The Neurology of mTOR. *Neuron* **84**, 275–291 (2014).
43. Zhao, J., Zhai, B., Gygi, S. P. & Goldberg, A. L. MTOR inhibition activates overall protein degradation by the ubiquitin proteasome system as well as by autophagy. *Proc Natl Acad Sci U S A* **112**, 15790–15797 (2015).
44. Chiang, S. Y. *et al.* Usp11 controls cortical neurogenesis and neuronal migration through Sox11 stabilization. *Sci Adv* **7**, 6093–6105 (2021).
45. Yan, Y. *et al.* X-linked ubiquitin-specific peptidase 11 increases tauopathy vulnerability in women. *Cell* **185**, 3913-3930.e19 (2022).
46. Buckanovich, R. J. & Darnell, R. B. The neuronal RNA binding protein Nova-1 recognizes specific RNA targets in vitro and in vivo. *Mol Cell Biol* **17**, 3194–3201 (1997).
47. Rivero, O. *et al.* Cadherin-13, a risk gene for ADHD and comorbid disorders, impacts GABAergic function in hippocampus and cognition. *Translational Psychiatry 2015 5:10* **5**, e655–e655 (2015).
48. Ciatto, C. *et al.* T-cadherin structures reveal a novel adhesive binding mechanism. *Nature Structural & Molecular Biology 2010 17:3* **17**, 339–347 (2010).
49. Hayano, Y. *et al.* The role of T-cadherin in axonal pathway formation in neocortical circuits. *Development* **141**, 4784–4793 (2014).
50. Hawi, Z. *et al.* The role of cadherin genes in five major psychiatric disorders: A literature update. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* **177**, 168–180 (2018).
51. Forero, A. *et al.* Serotonin (5-HT) neuron-specific inactivation of Cadherin-13 impacts 5-HT system formation and cognitive function. *Neuropharmacology* **168**, 108018 (2020).
52. Whiting, J. R. *et al.* Drainage-structuring of ancestral variation and a common functional pathway shape limited genomic convergence in natural high- and low-predation guppies. *PLoS Genet* **17**, e1009566 (2021).
53. Seghers, B. H. Schooling Behavior in the Guppy (Poecilia reticulata): An Evolutionary Response to Predation. *Evolution (N Y)* **28**, 489 (1974).
54. Huizinga, M., Ghalambor, C. K. & Reznick, D. N. The genetic and environmental basis of adaptive differences in shoaling behaviour among populations of Trinidadian guppies, Poecilia reticulata. *J Evol Biol* **22**, 1860–1866 (2009).
55. Herbert-Read, J. E. *et al.* How predation shapes the social interaction rules of shoaling fish. *Proceedings of the Royal Society B: Biological Sciences* **284**, 20171126 (2017).
56. Strandburg-Peshkin, A. *et al.* Visual sensory networks and effective information transfer in animal groups. *Current biology* **23**, R711 (2013).
57. Conradt, L. & Roper, T. J. Conflicts of interest and the evolution of decision sharing. *Philosophical Transactions of the Royal Society B: Biological Sciences* **364**, 807–819 (2008).
58. Ioannou, C. C., Singh, M. & Couzin, I. D. Potential Leaders Trade Off Goal-Oriented and Socially Oriented Behavior in Mobile Animal Groups. *Am Nat* **186**, 284–293 (2015).
59. Bshary, R., Gingins, S. & Vail, A. L. Social cognition in fishes. *Trends Cogn Sci* **18**, 465–471 (2014).
60. Sinha, S. *et al.* Behavior-related gene regulatory networks: A new level of organization in the brain. *Proc Natl Acad Sci U S A* **117**, 23270–23279 (2020).

872 61. Hoekstra, H. E. & Coyne, J. A. THE LOCUS OF EVOLUTION: EVO DEVO AND THE
873 GENETICS OF ADAPTATION. *Evolution (N Y)* **61**, 995–1016 (2007).
874 62. Carroll, S. B. Evo-Devo and an Expanding Evolutionary Synthesis: A Genetic Theory of
875 Morphological Evolution. *Cell* **134**, 25–36 (2008).
876 63. Bloch, N. I. *et al.* Early neurogenomic response associated with variation in guppy female
877 mate preference. *Nature Ecology & Evolution 2018 2:11* **2**, 1772–1781 (2018).
878 64. Young, R. L. *et al.* Conserved transcriptomic profiles underpin monogamy across
879 vertebrates. *Proc Natl Acad Sci U S A* **116**, 1331–1336 (2019).
880 65. Rittschof, C. C. *et al.* Neuromolecular responses to social challenge: Common
881 mechanisms across mouse, stickleback fish, and honey bee. *Proc Natl Acad Sci U S A*
882 **111**, 17929–17934 (2014).
883 66. Pélabon, C. *et al.* The effects of sexual selection on life-history traits: an experimental
884 study on guppies. *J Evol Biol* **27**, 404–416 (2014).
885 67. Pérez-Escudero, A., Vicente-Page, J., Hinz, R. C., Arganda, S. & De Polavieja, G. G.
886 idTracker: tracking individuals in a group by automatic identification of unmarked
887 animals. *Nature Methods 2014 11:7* **11**, 743–748 (2014).
888 68. Kotrschal, A. *et al.* Brain size does not impact shoaling dynamics in unfamiliar groups of
889 guppies (Poecilia reticulata). *Behavioural Processes* **147**, 13–20 (2018).
890 69. Corral-López, A., Romensky, M., Kotrschal, A., Buechel, S. D. & Kolm, N. Brain size
891 affects responsiveness in mating behaviour to variation in predation pressure and sex ratio.
892 *J Evol Biol* **33**, (2020).
893 70. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation
894 for Statistical Computing, Vienna, Austria. https://www.R-project.org/ (2021).
895 71. Posit team. RStudio: Integrated Development Environment for R. Posit Software, PBC,
896 Boston, MA. URL http://www.posit.co/ (2023).
897 72. Wickham, H. *et al.* Welcome to the Tidyverse. *J Open Source Softw* **4**, 1686 (2019).
898 73. Bates, D., Sarkar, D., Bates, M. D. & Matrix, L. The lme4 Package. *R package version* **2**,
899 74 (2007).
900 74. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest Package: Tests in
901 Linear Mixed Effects Models. *J Stat Softw* **82**, 1–26 (2017).
902 75. de Villemereuil, P. On the relevance of Bayesian statistics and MCMC for animal models.
903 *Animal Breeding Genetics* **136**, 339–340 (2019).
904 76. Bürkner, P. C. Advanced Bayesian multilevel modeling with the R package brms. *R J* **10**,
905 395–411 (2018).
906 77. Bürkner, P. C. brms: An R package for Bayesian multilevel models using Stan. *J Stat*
907 *Softw* **80**, 1–28 (2017).
908 78. Carpenter, B. *et al.* Stan: A probabilistic programming language. *J Stat Softw* **76**, (2017).
909 79. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina
910 sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
911 80. Künstner, A. *et al.* The Genome of the Trinidadian Guppy, Poecilia reticulata, and
912 Variation in the Guanapo Population. *PLoS One* **11**, e0169087 (2016).
913 81. Pertea, M., Kim, D., Pertea, G. M., Leek, J. T. & Salzberg, S. L. Transcript-level
914 expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown.
915 *Nature Protocols 2016 11:9* **11**, 1650–1667 (2016).
916 82. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**,
917 2078–2079 (2009).

83. Turner, S. D. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *J Open Source Softw* **3**, 731 (2018).
84. Mi, H., Muruganujan, A., Casagrande, J. T. & Thomas, P. D. Large-scale gene function analysis with the PANTHER classification system. *Nature Protocols 2013 8:8* **8**, 1551–1566 (2013).
85. Li, B. & Dewey, C. N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 1–16 (2011).
86. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 1–21 (2014).
87. Amar, D., Safer, H. & Shamir, R. Dissection of Regulatory Networks that Are Altered in Disease via Differential Co-expression. *PLoS Comput Biol* **9**, e1002955 (2013).
88. Fukushima, A. DiffCorr: An R package to analyze and visualize differential correlations in biological networks. *Gene* **518**, 209–214 (2013).
89. Rezwani, M., Pourfathollah, A. A. & Noorbakhsh, F. rbioapi: user-friendly R interface to biologic web services' API. *Bioinformatics* **38**, 2952–2953 (2022).
90. Reimand, J. *et al.* g:Profiler—a web server for functional interpretation of gene lists (2016 update). *Nucleic Acids Res* **44**, W83–W89 (2016).
91. Karaca, E. *et al.* Genes that Affect Brain Structure and Function Identified by Rare Variant Analyses of Mendelian Neurologic Disease. *Neuron* **88**, 499–513 (2015).
92. Takeichi, M. The cadherin superfamily in neuronal connections and interactions. *Nature Reviews Neuroscience 2006 8:1* **8**, 11–20 (2006).

Figures and Tables

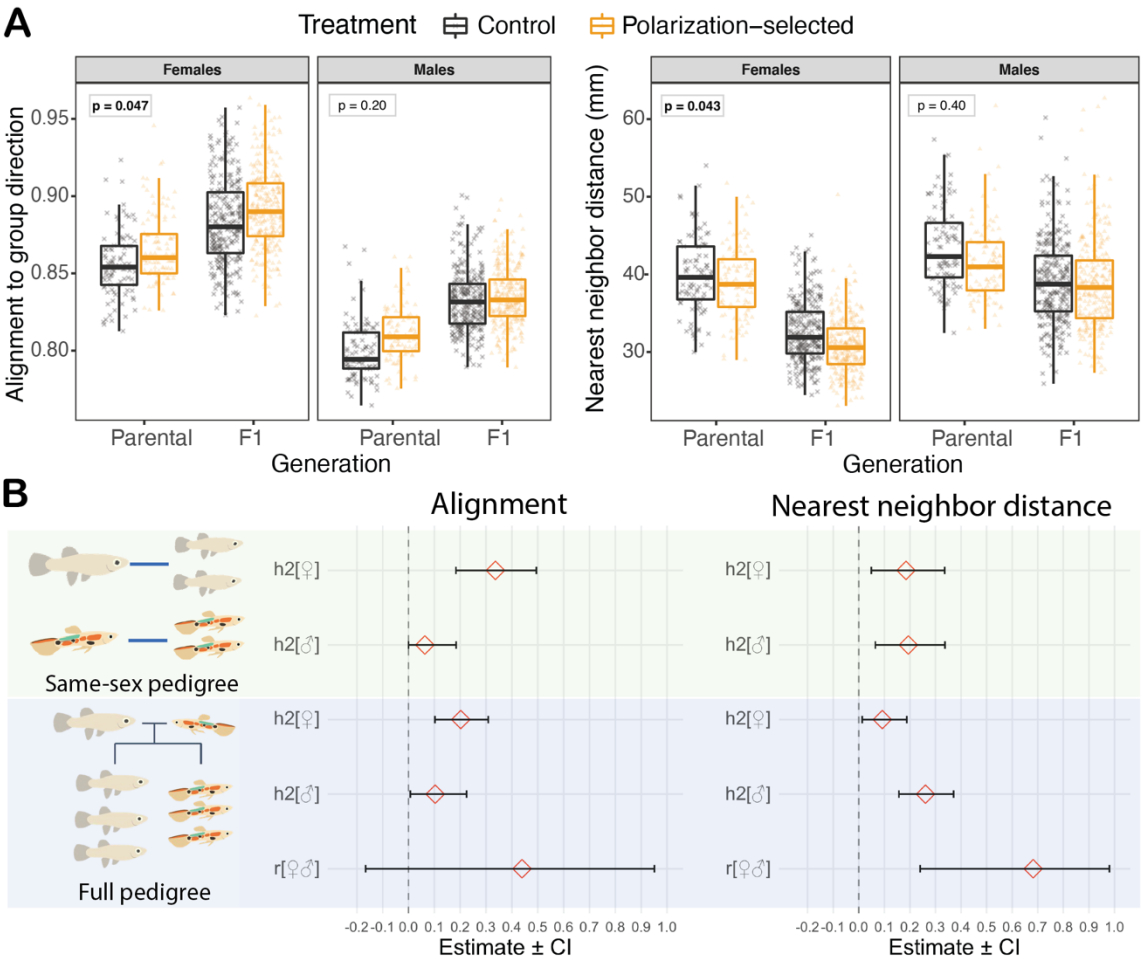


Figure 1. Heritability of sociability in guppies. (A) Female, but not male, guppies from polarization-selected lines (orange) presented higher alignment to the group direction (left panel) and shorter distances to their nearest neighbor (higher alignment; right panel) than guppies from control lines (gray) when swimming with unfamiliar same-sex conspecifics (see Supplementary Table 1-2). For all boxplots, horizontal lines indicate medians, boxes indicate the interquartile range, and whiskers indicate all points within 1.5 times the interquartile range. Boxes in top left position of each facet indicate p-values ($p < 0.05$ in bold) for statistical contrasts by sex in Linear Mixed Model comparing alignment and attraction between selection line treatments (see Supplementary Tables 1-2). **(B)** Animal models using same-sex pedigrees and full pedigrees with alignment and attraction (nearest neighbor distance) phenotypes in 195 families of polarization-selected and control guppy lines indicated a moderate heritability in female guppies of both biologically-relevant aspects of sociability measured, alignment (left) and attraction (right). In males, we found moderate heritability in attraction, but credible intervals in alignment estimates overlapped with 0 suggesting low heritability of this sociability aspect. Our full pedigree animal models provided large credible intervals for male-female correlations in sociability, with estimates overlapping 0 in alignment, but a positive correlation between sexes in attraction (see Supplementary Tables 4-5.

974 **Table 1. Characterization of genes associated with sociability in guppies**
975

SNP location	Gene ID Ensembl	Gene name	Described cognitive function of homologs	References
Chr 7: 30748139	00000004543	*usp11*	Control of cortical neurogenesis and neuronal migration. Mutations of the gene have been associated with neurological disorders.	Chiang et al, 2021[44] Karaca et al. 2015[91]
Chr 13: 31383940	00000018946	*Novel gene (RNA-binding protein Nova-1-like)*	Neuronal RNA binding protein associated with motor function.	Buckanovich and Darnell 1997[46]
Chr 14: 4286109	00000009725	*supt6h*	Substrate of mTOR, a signaling pathway associated with brain function and neurodegenerative disorders.	Lipton and Sahin 2015[42] Zhao et al. 2015[43]
Chr 18: 4286109	00000014318	*Novel gene*	-	-
Chr 19: 3032099	00000019822	*cdh13*	Modulation of brain activity through GABAergic function. Organization of neuronal circuits.	Rivero et al 2015[47] Takeichi 2006[92]

976
977
978
979
980
981
982
983
984
985
986
987

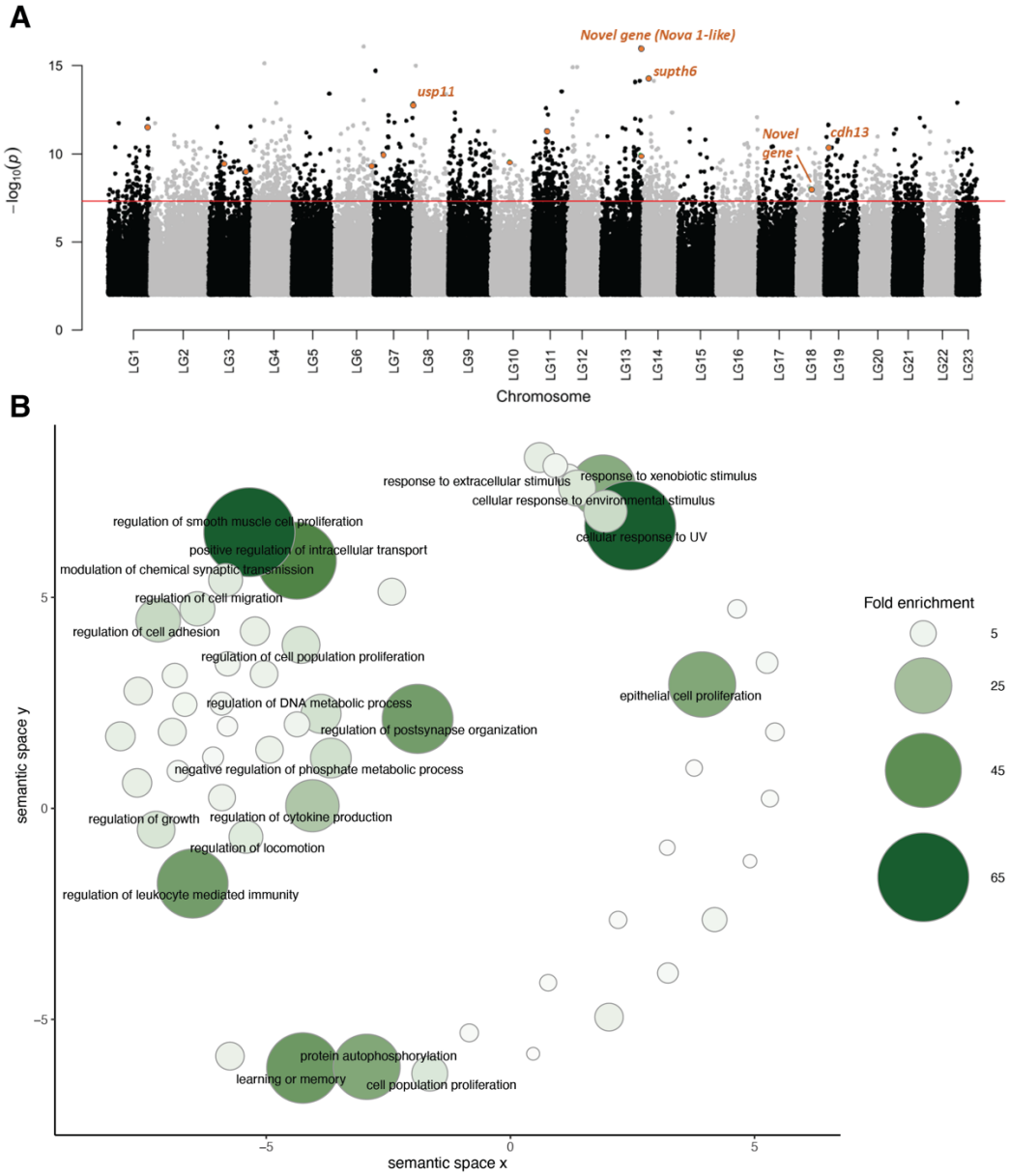


Figure 2. Genetic basis of sociability in the guppy. (A) Manhattan plot of -log10(p) values across the guppy genome resulting from a Fisher's exact test comparing allele frequency differences between high and low sociability female guppies. We merged pooled-DNA sequences of three independent replicates and found 819 SNPs significantly different (above genome-wide threshold highlighted in red), while a highly stringent analyses of consistent allele frequency differences across our three independent replicates (cmh-test; see methods) identified 13 SNPs (five of them within genes) associated with sociability in the species (gene names and SNP location in the genome highlighted in orange). SNPs with -log10(p) values < 2 are omitted. **(B)** Clustering of statistically significant overrepresented Gene Ontology annotations for biological processes associated with differences between high and low sociability in guppies. Point size and color provide information on fold enrichment value from the statistical overrepresentation test performed in PANTHER[84] (see methods). Terms with fold enrichment lower than 8 are represented but not described in text. Axes have no intrinsic meaning and are based on multidimensional scaling which clustered terms based on semantic similarities[74].

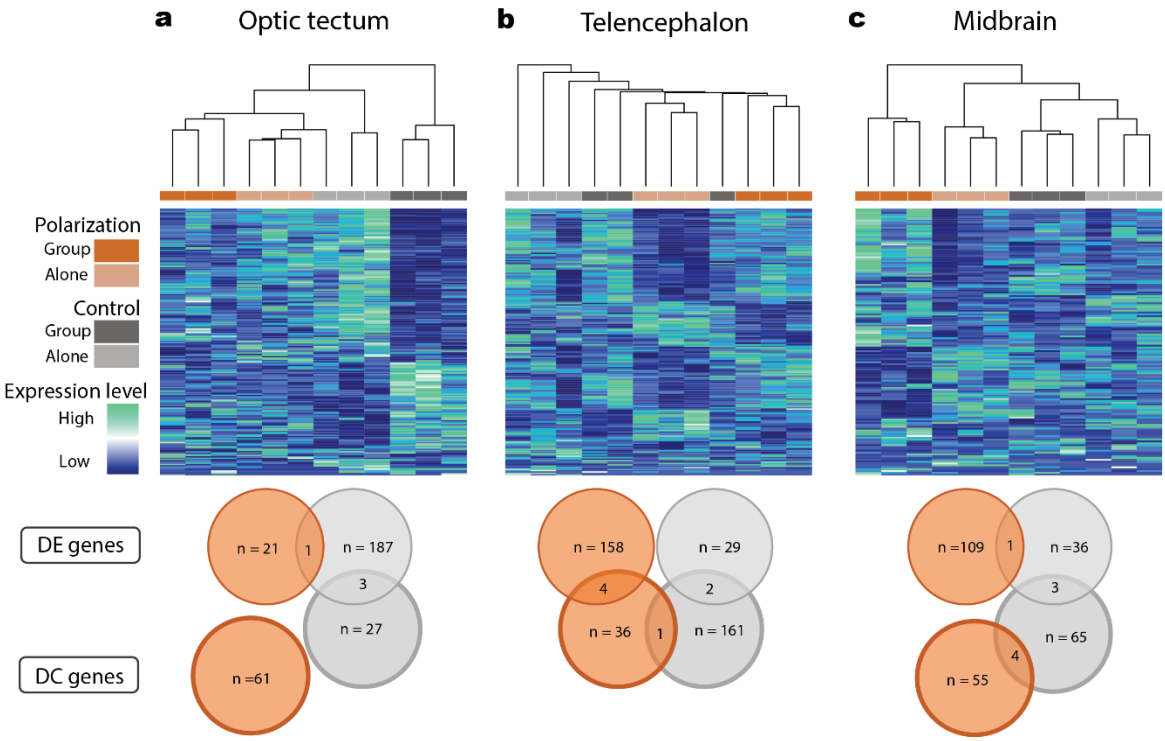


Figure 3. Neurogenomic response of schooling in guppies. Hierarchical clustering and relative expression level for all differentially expressed genes between Alone and Group treatments in the (a) optic tectum, (b) the telencephalon and (c) the midbrain. Differentially expressed genes were identified separately in polarization and control line samples. Clustering, based on Euclidian distance, represents transcriptional similarity across all samples, with bootstrap values (1000 replicates) shown at nodes. Venn diagrams summarize the total number of differentially expressed (DE) genes and differentialy coexpressed (DC) gene pairs in each tissue for polarization-selected and control lines.

Figure 4. Functional characterization of genes of interest across experiments. Visualization of functional overlap based on Gene Ontology annotations between genes of interest highlighted in strongly differentiated experimental setups evaluating social interactions of female guppies following experimental evolution for higher polarization: 1- genomic analyses of DNA comparing Pool-seq of high and low sociability female guppies (Left Column); 2- transcriptomic analyses evaluating differentially expressed genes in key brain regions of polarization-selected lines of female guppies exposed to two different social contexts, swimming alone or with a group of conspecifics (TEL: telencephalon, MBR: midbrain, OT: optic tectum; Columns 2-4). Shades of green indicate fold enrichment from our statistical overrepresentation tests performed to gene lists obtained from each experiment.

Supplementary Figures for:

Functional convergence of genomic and transcriptomic genetic architecture underlying sociability in a live-bearing fish

Alberto Corral-Lopez*, Natasha I. Bloch, Wouter van der Bijl, Maria Cortazar-Chinarro, Alexander Szorkovszky, Alexander Kotrschal, Iulia Darolti, Severine D. Buechel, Maksym Romenskyy, Niclas Kolm, Judith E. Mank

*corresponding author: alberto.corral@ebc.uu.se

Data sets and code available at 10.6084/m9.figshare.23805702
Sequence data available at NCBI repository (accession codes PRJNA994132 and PRJNA504011)

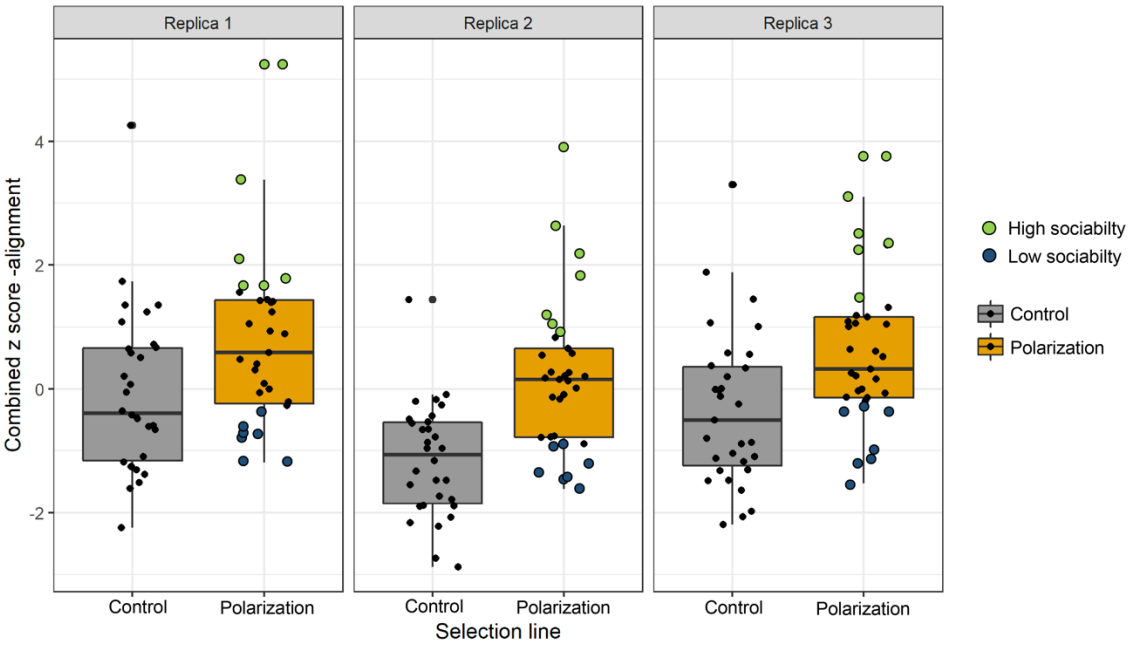


Figure S1. Mean score for alignment to group direction in females for each of the 195 families measured when evaluating the heritability of sociability by means of open field tests in which guppies from three independently replicated polarization-selected (orange) and control (gray) lines were exposed to a group of other seven conspecifics. We obtained combined Z scores by standardizing independently alignment scores obtained in maternal and offspring generations and by adding the score of the mother to the mean score of all female offspring measure in each family. For all boxplots, horizontal lines indicate medians, boxes indicate the interquartile range, and whiskers indicate all points within 1.5 times the interquartile range. Colored circles indicate families that were further used for genomic analyses.

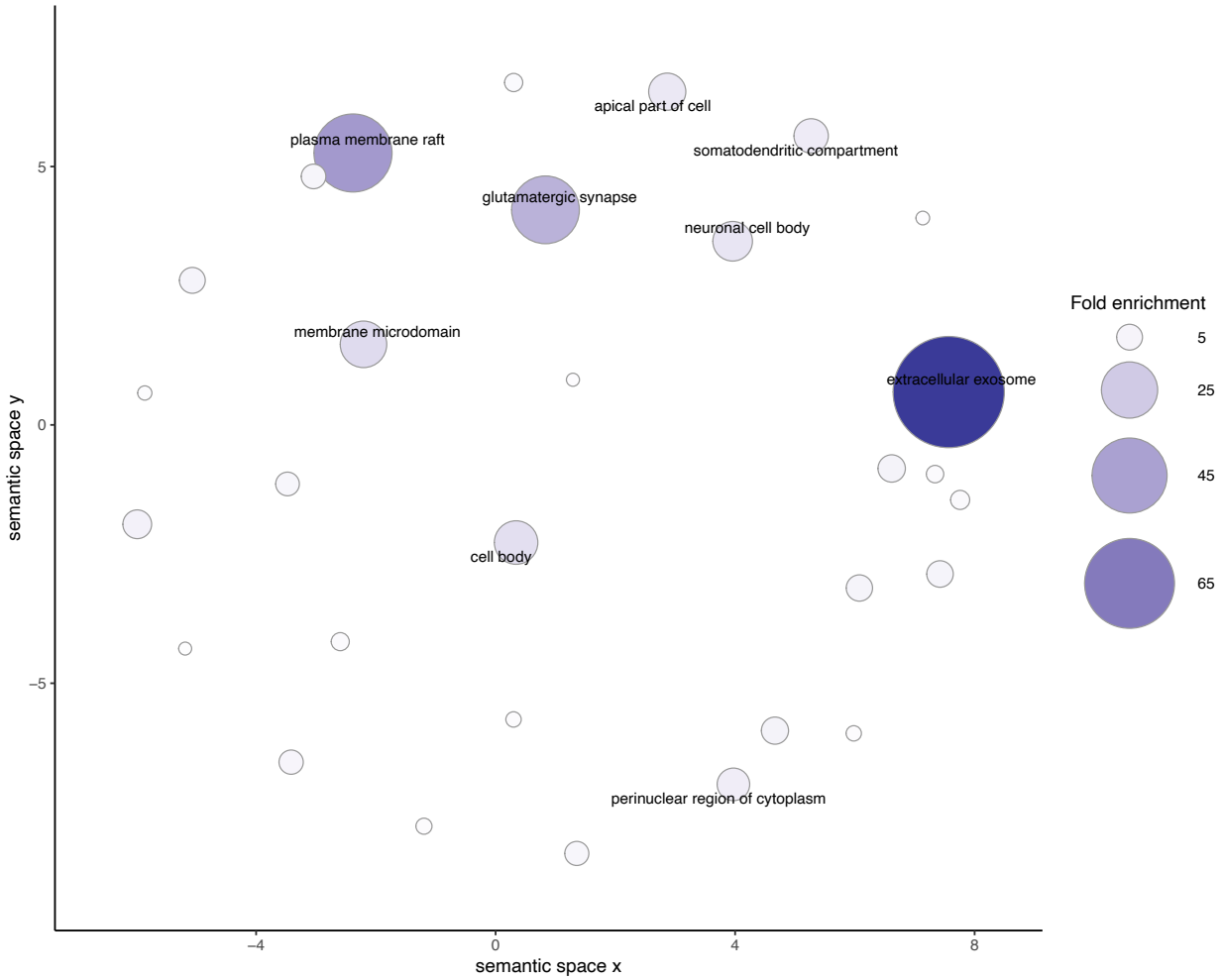


Figure S2. Clustering of statistically significant overrepresented Gene Ontology annotations for cellular components associated to differences between high and low sociability in guppies. Point size and color provide information on fold enrichment value from the statistical overrepresentation test performed in PANTHER[1]. Terms with fold enrichment lower than eight are represented but not described in text. Axes have no intrinsic meaning and are based on multidimensional scaling which cluster terms based on semantic similarities[2].

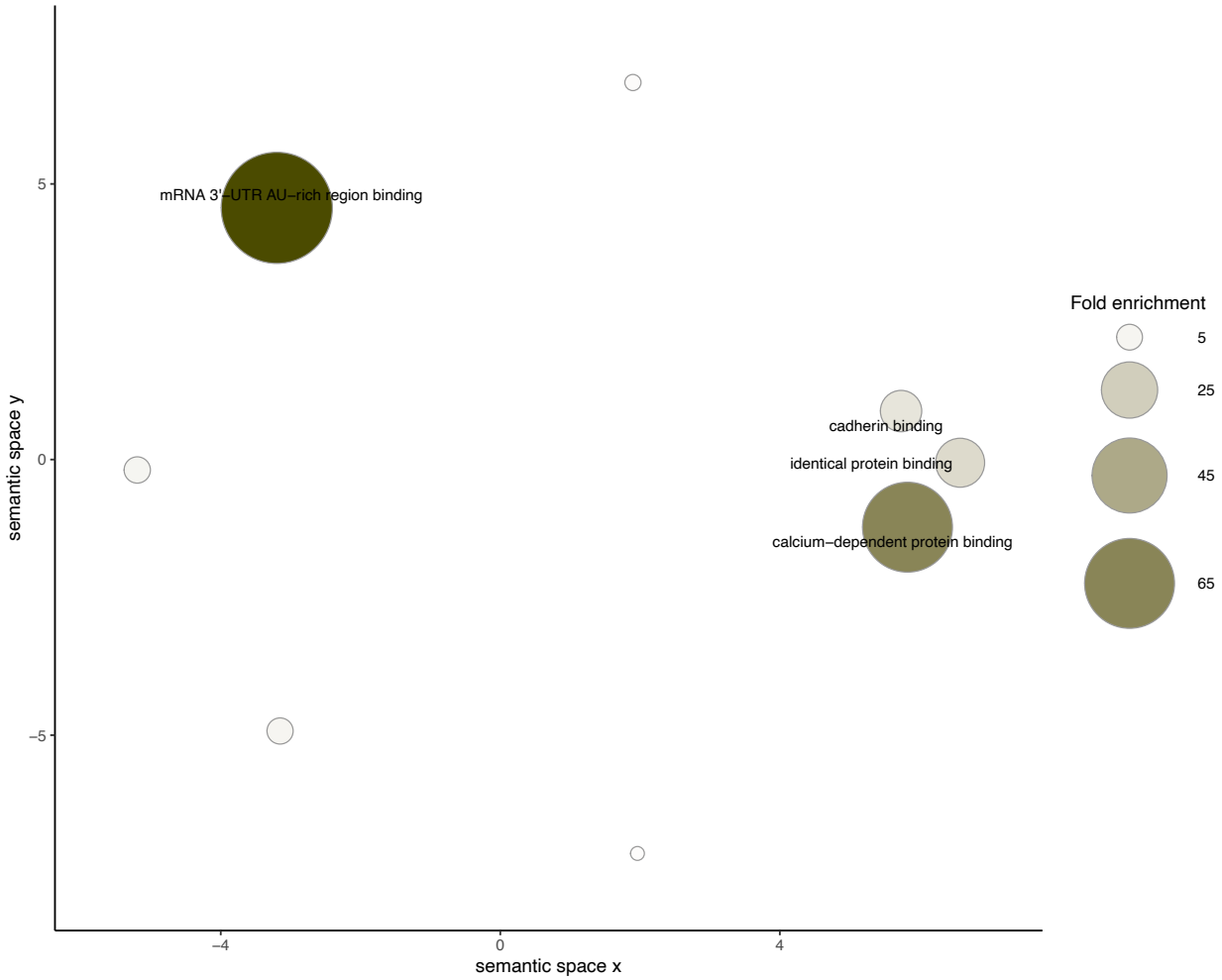


Figure S3. Clustering of statistically significant overrepresented Gene Ontology annotations for molecular functions associated to differences between high and low sociability in guppies. Point size and color provide information on fold enrichment value from the statistical overrepresentation test performed in PANTHER[1]. Terms with fold enrichment lower than eight are represented but not described in text. Axes have no intrinsic meaning and are based on multidimensional scaling which cluster terms based on semantic similarities[2].

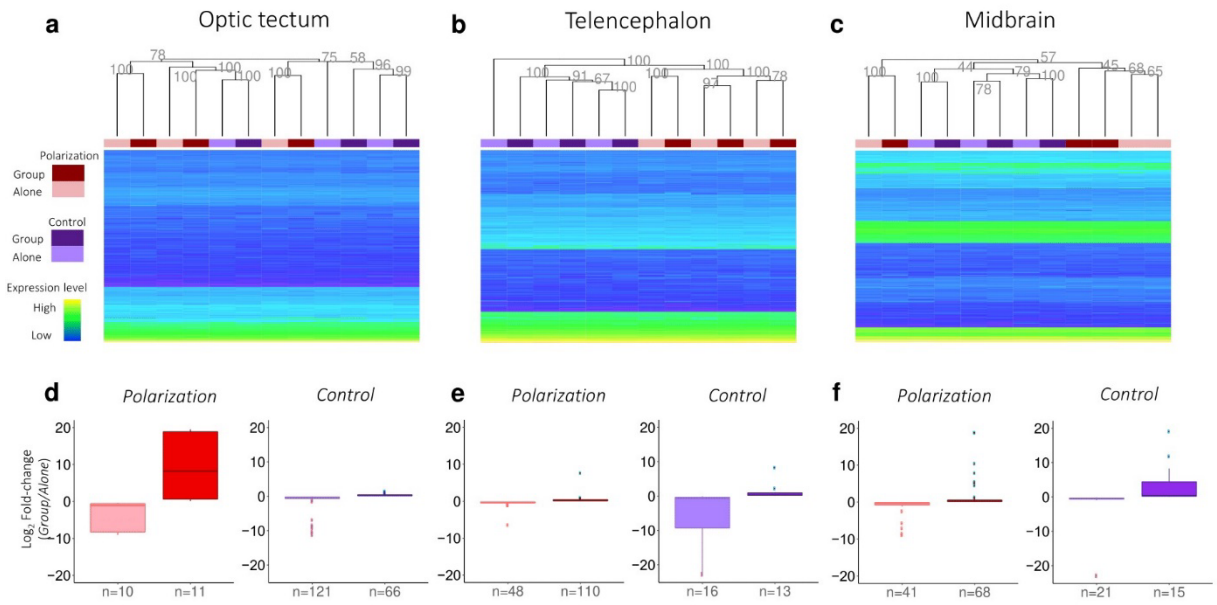


Figure S4. Top panels: Hierarchical clustering of gene expression across all significantly expressed genes in the optic tectum (a), telencephalon (b) and midbrain (c) for polarization-selected selection and control lines. Values on top of nodes correspond to Approximately Unbiased bootstrap values. Heatmap represents expression of 500 randomly selected genes. **Bottom panels:** Boxplots of average log2 fold-change for all genes differentially expressed genes between the *Alone* and the *Group* conditions in polarization-selected and control lines for optic tectum (d), telencephalon (e) and midbrain (f). Boxes correspond to 25th - 75th percentiles. Sample sizes on the x-axis indicate number of genes with relative up- or down-regulation.

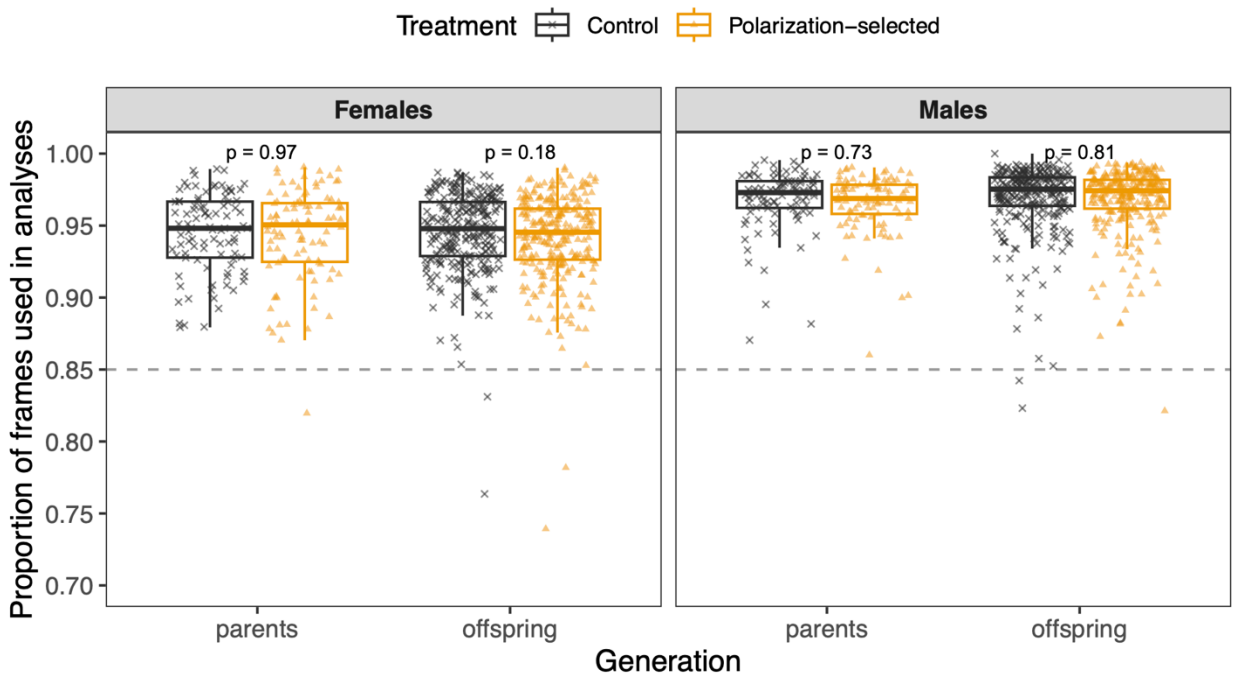


Figure S5. Proportion of frames obtained from idTracker software in tracking from videos recorded for groups of seven wild-type guppies and one individual from either polarization-selected or control selection lines. Tracking data was used next for the quantification of alignment, attraction and speed in parents and six offspring of 195 families. We found no significant differences for the proportion of frames used for videos used for polarization-selected and control lines in a Linear Mixed Model that included generation, sex and selection line as fixed factors (LMM_{frames}: line: F= 0.82; df = 1464; p = 0.37). Horizontal lines indicate medians, boxes indicate the interquartile range, and whiskers indicate all points within 1.5 times the interquartile range. P-values in top position of each comparison indicate values for statistical contrasts of the model by sex and generation. Consistency in identity assigned by idTracker was verified in trials with less than 85% frames tracked (dashed line).

Supplementary References

1 - Mi, H., Muruganujan, A., Casagrande, J. T. & Thomas, P. D. Large-scale gene function analysis with the PANTHER classification system. *Nature Protocols 2013 8:8* **8**, 1551–1566 (2013).

2 - Supek, F., Bošnjak, M., Škunca, N. & Šmuc, T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* **6**, (2011).

Supplementary Tables for:

Functional convergence of genomic and transcriptomic genetic architecture underlying sociability in a live-bearing fish

Alberto Corral-Lopez*, Natasha I. Bloch, Wouter van der Bijl, Maria Cortazar-Chinarro, Alexander Szorkovszky, Alexander Kotrschal, Iulia Darolti, Severine D. Buechel, Maksym Romenskyy, Niclas Kolm, Judith E. Mank

*corresponding author: alberto.corral@ebc.uu.se

Data sets and code available at 10.6084/m9.figshare.23805702
Sequence data available at NCBI repository (accession codes PRJNA994132 and PRJNA504011)

Table S1. Statistical results for a Linear Mixed Model comparing the alignment and attraction of polarization-selected and control female guppies when swimming with groups of same-sex non-kin conspecifics.

Alignment to group					
Predictors	*Estimates*	*CI*	*Statistic*	*p*	*df*
(Intercept)	0.19	0.13 – 0.24	6.78	**<0.001**	1385.81
Selection line [polarization]	0.01	0.00 – 0.01	2.27	**0.047**	9.68
Sex [males]	-0.01	-0.01 – -0.00	-3.02	**0.003**	690.08
Generation [parental]	-0.02	-0.02 – -0.01	-10.13	**<0.001**	1141.24
Average alignment of group	0.79	0.72 – 0.86	22.71	**<0.001**	1436.72
Average speed of group	0.00	0.00 – 0.00	0.20	0.842	1429.55
Selection line [polarization] * Sex [males]	0.00	-0.00 – 0.00	0.20	0.842	1429.55
ICC	0.12				
N_{line}	2				
N_{rep}	3				
N_{testnr}	8				
N_{sex}	2				
N_{family}	195				
Observations	1486				
Marginal R^2 / Conditional R^2	0.657 / 0.697				

Nearest neighbour distance (attraction)					
Predictors	*Estimates*	*CI*	*Statistic*	*p*	*df*
(Intercept)	2.64	2.56 – 2.71	65.62	**<0.001**	565.12
Selection line [polarization]	-0.04	-0.08 – -0.00	-2.33	**0.044**	9.34
Sex [males]	0.01	-0.02 – 0.04	0.55	0.583	448.53
Generation [parental]	0.12	0.10 – 0.14	11.31	**<0.001**	984.50
Average alignment of group	0.02	0.02 – 0.02	27.31	**<0.001**	1408.27

Average speed of group	0.00	0.00 – 0.00	3.74	**<0.001**	566.82
Selection line [polarization] * Sex [males]	0.03	-0.02 – 0.07	1.24	0.217	371.59
ICC	0.13				
N_{line}	2				
N_{rep}	3				
N_{testnr}	8				
N_{sex}	2				
N_{family}	195				
Observations	1485				
Marginal R^2 / Conditional R^2	0.471 / 0.538				

Table S2. Independent contrasts for comparisons between polarization-selected (P) and control (C) guppies in their alignment and nearest neighbor distance (attraction) when swimming with groups of same-sex non-kin conspecifics.

Alignment

contrast	*sex*	*estimate*	*SE*	*df*	*t.ratio*	*p.value*
C - P	females	-0.006	0.002	9.68	-2.269	**0.047**
C - P	males	-0.003	0.002	9.56	1.378	0.20

Attraction

contrast	*treatment*	*estimate*	*SE*	*df*	*t.ratio*	*p.value*
C - P	females	0.042	0.001	9.34	2.331	**0.043**
C - P	males	0.015	0.002	9.26	0.883	0.40

Table S3. Mean body size of guppies used to estimate the heritability of alignment and attraction at the time of testing across sexes and generations. Values are provided in number of pixels obtained from idTracker data using a custom script implemented in Matlab (v2020a) that extracts body size data accounting for changes in apparent size between the middle and edge of the experimental arena used for video recordings.

Generation	Selection line	Sex	Mean body size	SE
Parents	Control	Females	410	47
		Males	192	29
	Polarization-selected	Females	416	71
		Males	189	37
Offspring	Control	Females	247	55
		Males	158	33
	Polarization-selected	Females	251	60
		Males	160	26

Table S4. Results from animal models only accounting for same-sex pedigree relationships. We obtained heritability estimates by taking the ratio of the additive genetic variance to the total phenotypic variance (additive genetic variance + group variance + population variance + residual variance) in each independent model.

ALIGNMENT

FEMALES

Parameter group	Parameter	Estimate	Est.Error	Q2.5	Q97.5
Fixed effects	Intercept	0.504	0.074	0.359	0.648
Fixed effects	linesocial	0.157	0.102	-0.04	0.355
Fixed effects	generationparents	-0.277	0.058	-0.39	-0.163
Fixed effects	speed_std	0.396	0.023	0.35	0.441
Additive genetic variance	Intercept	0.162	0.042	0.086	0.25
Group	Intercept	0.005	0.006	0	0.02
Population	Intercept	0.01	0.027	0	0.064
Residual variance	NA	0.304	0.037	0.234	0.379
Heritability		0.336	0.079	0.183	0.494

MALES

Parameter group	Parameter	Estimate	Est.Error	Q2.5	Q97.5
Fixed effects	Intercept	-0.452	0.125	-0.694	-0.213
Fixed effects	linesocial	0.093	0.172	-0.231	0.425
Fixed effects	generationparents	-0.502	0.071	-0.641	-0.362
Fixed effects	speed_std	0.298	0.039	0.221	0.376
Additive genetic variance	Intercept	0.031	0.026	0.000	0.091
Group	Intercept	0.04	0.016	0.013	0.074
Population	Intercept	0.039	0.091	0.001	0.207
Residual variance	NA	0.390	0.033	0.323	0.453
Heritability		0.063	0.052	0.000	0.184

ATTRACTION

FEMALES

Parameter group	Parameter	Estimate	Est.Error	Q2.5	Q97.5
Fixed effects	Intercept	-0.614	0.099	-0.811	-0.424
Fixed effects	linesocial	-0.088	0.138	-0.35	0.179
Fixed effects	generationparents	0.251	0.070	0.113	0.390
Fixed effects	speed_std	-0.109	0.024	-0.157	-0.062
Additive genetic variance	Intercept	0.094	0.038	0.025	0.174
Group	Intercept	0.038	0.015	0.011	0.071
Population	Intercept	0.023	0.061	0.000	0.133
Residual variance	NA	0.355	0.039	0.280	0.435
Heritability		0.185	0.072	0.049	0.336

MALES

Parameter group	Parameter	Estimate	Est.Error	Q2.5	Q97.5
Fixed effects	Intercept	0.476	0.134	0.212	0.746
Fixed effects	linesocial	-0.025	0.181	-0.389	0.345
Fixed effects	generationparents	0.614	0.088	0.440	0.785
Fixed effects	speed_std	0.054	0.046	-0.037	0.145
Additive genetic variance	Intercept	0.137	0.049	0.047	0.242
Group	Intercept	0.075	0.023	0.036	0.126
Population	Intercept	0.046	0.097	0.000	0.250
Residual variance	NA	0.454	0.050	0.359	0.556
Heritability		0.194	0.069	0.065	0.337

Table S5 Results from animal models with full pedigree relationships. We obtained heritability estimates for each sex by taking the ratio of the additive genetic variance of each sex to the total phenotypic variance (additive genetic variance + group variance + population variance + residual variance) in each independent model.

ALIGNMENT

Parameter group	Parameter	Estimate	Est.Error	Q2.5	Q97.5
Fixed effects	Intercept	0.52	0.102	0.331	0.718
Fixed effects	sexm	-0.961	0.063	-1.08	-0.838
Fixed effects	linesocial	0.163	0.128	-0.097	0.419
Fixed effects	speed_std	0.36	0.021	0.319	0.4
Fixed effects	sexm:linesocial	-0.075	0.074	-0.22	0.07
Fixed effects	sexf:generationparents	-0.31	0.066	-0.444	-0.181
Fixed effects	sexm:generationparents	-0.479	0.064	-0.604	-0.352
Additive genetic variance	sexf	0.105	0.029	0.051	0.166
Additive genetic variance	sexm	0.048	0.027	0.003	0.105
Additive genetic variance	cross (sexf x sexm)	0.583	0.270	0.021	0.974
Group	Intercept	0.025	0.009	0.009	0.045
Population	Intercept	0.023	0.051	0.001	0.13
Residual variance	NA	0.365	0.026	0.314	0.415
	♀ Heritability	0.202	0.052	0.102	0.308
	♂ Heritability	0.103	0.057	0.007	0.224
	Δ Heritability (♀ - ♂)	0.098	0.053	0.202	-0.005

ATTRACTION

Parameter group	Parameter	Estimate	Est.Error	Q2.5	Q97.5
Fixed effects	Intercept	-0.627	0.115	-0.847	-0.391
Fixed effects	sexm	1.09	0.075	0.942	1.23
Fixed effects	linesocial	-0.095	0.157	-0.425	0.219
Fixed effects	speed_std	-0.067	0.022	-0.109	-0.024
Fixed effects	sexm:linesocial	0.072	0.082	-0.091	0.235
Fixed effects	sexf:generationparents	0.283	0.079	0.129	0.438
Fixed effects	sexm:generationparents	0.568	0.080	0.411	0.722
Additive genetic variance	sexf	0.050	0.025	0.007	0.104
Additive genetic variance	sexm	0.176	0.041	0.102	0.263
Additive genetic variance	sexm	0.438	0.304	-0.168	0.950
Group	Intercept	0.059	0.013	0.035	0.087
Population	Intercept	0.032	0.056	0.002	0.172
Residual variance	NA	0.406	0.028	0.351	0.460
	♀ Heritability	0.092	0.044	0.013	0.188
	♂ Heritability	0.261	0.055	0.157	0.371
	Δ Heritability (♀ - ♂)	-0.169	0.051	-0.267	-0.069

Table S6. Differentially expressed genes in analyses of the genomic response in midbrain, optic tectum and telencephalon in polarization-selected and control lines of female guppies when swimming with a group of conspecifics versus when swimming alone.

MIDBRAIN

CONTROL LINES (n=36)

Transcript	Gene BLAST	Average expression	Fold Change	P-value (FDR)
rna32553|gene=LOC103477857	PLCL2	131.212	27.132	<0.001
MSTRG.16482.1|gene=MSTRG.16482	MORN4	37.361	-22.976	<0.001
MSTRG.17365.1|gene=MSTRG.17365	RPGP2	32.181	-22.742	<0.001
rna23930|gene=ubap2l	ubap2l	944.324	-0.346	<0.001
rna14333|gene=trim33	trim33	41.183	19.099	<0.001
rna12705|gene=LOC103466075	MAGI2	1073.054	-0.398	<0.001
rna14220|gene=LOC103466954	FLNA	259.466	8.280	<0.001
rna43613|gene=LOC103458483	SYNE1	7853.703	0.243	<0.001
rna46407|gene=lrfn5	lrfn5	1152.716	-0.665	<0.001
rna12704|gene=LOC103466075	MAGI2	1284.870	0.358	0.001
MSTRG.18176.1|gene=MSTRG.18176	KMCP1	206.043	-0.800	0.001
rna32800|gene=u2af2	u2af2	1910.842	-0.346	0.001
MSTRG.16838.1|gene=MSTRG.16838	K1841	386.297	0.320	0.002
MSTRG.4096.1|gene=MSTRG.4096	AN13C	167.036	1.072	0.002
rna9959|gene=kcnab2	kcnab2	2321.886	-0.139	0.003
MSTRG.14391.4|gene=MSTRG.14391	KMT2A	2251.545	0.175	0.004
rna6890|gene=LOC103462605	IGF1R	151.381	5.464	0.004
MSTRG.16227.1|gene=MSTRG.16227	RIBA2	144.642	-0.623	0.004
rna27802|gene=fndc3a	fndc3a	639.741	-0.517	0.004
rna42558|gene=itpkb	itpkb	253.080	-0.714	0.006
MSTRG.6879.2|gene=MSTRG.6879	AEBP2	492.647	-0.492	0.009
rna15437|gene=LOC103467579	IQEC1	272.493	-0.544	0.010
MSTRG.10425.1|gene=MSTRG.10425	METK2	2522.539	-0.149	0.010
rna39037|gene=tanc2	tanc2	352.805	11.879	0.011
MSTRG.17415.2|gene=MSTRG.17415	FLOT1	1005.390	-0.284	0.018
MSTRG.3778.3|gene=MSTRG.3778	MAP1B	10822.870	0.090	0.022
rna19452|gene=LOC103470078	SCAM1	1692.812	0.214	0.027
rna30679|gene=LOC103476732	ANK3	6992.816	0.149	0.027
rna42097|gene=LOC103457647	NA	245.652	-0.955	0.027
MSTRG.7457.1|gene=MSTRG.7457	FLNA	1583.394	-0.378	0.033
MSTRG.7577.1|gene=MSTRG.7577	F19A5	1752.482	0.375	0.035
rna1276|gene=acsl1	acsl1	1052.049	-0.289	0.037
rna46838|gene=LOC103460240	FBL10	99.185	-0.667	0.038
rna18681|gene=LOC103469777	SMAD2	106.022	-0.567	0.042

rna14493|gene=wdr6	wdr6	151.669	0.864	0.048
rna1738|gene=keap1	keap1	2161.342	-0.225	0.049

POLARIZATION-SELECTED LINES (n=109)

Transcript	Gene BLAST	Average expression	Fold Change	P-value (FDR)
rna7061|gene=LOC103462879	PPR3E	144.810	10.467	< 0.001
rna9051|gene=LOC103463550	TLE4	90.600	-24.816	< 0.001
rna24376|gene=scn1b	scn1b	86.552	24.017	< 0.001
rna44020|gene=srrm1	srrm1	13.933	21.607	< 0.001
MSTRG.24305.1|gene=MSTRG.24305	NA	2510.567	0.748	< 0.001
rna22214|gene=egr1	egr1	30.831	18.839	< 0.001
rna22213|gene=egr1	egr1	30.831	18.839	< 0.001
rna40013|gene=LOC103482261	CPNE3	115.525	-1.047	< 0.001
rna45945|gene=syn3	syn3	2021.198	0.623	< 0.001
MSTRG.17634.6|gene=MSTRG.17634	CHIO	726.382	0.310	< 0.001
MSTRG.11465.4|gene=MSTRG.11465	LAMP2	1545.620	0.231	< 0.001
MSTRG.13269.1|gene=MSTRG.13269	FBCD1	2483.148	0.135	< 0.001
MSTRG.24306.3|gene=MSTRG.24306	SYN3	3100.717	1.061	< 0.001
rna26833|gene=LOC103474460	NCPR	105.368	5.135	0.001
MSTRG.3790.1|gene=MSTRG.3790	NA	6932.462	0.120	0.001
rna46319|gene=rps14	rps14	6939.074	-0.112	0.001
MSTRG.25027.1|gene=MSTRG.25027	NA	440.681	0.287	0.001
rna29486|gene=ubb	ubb	1424.005	-0.204	0.001
MSTRG.3031.1|gene=MSTRG.3031	NRK2	907.778	-0.234	0.001
rna35365|gene=LOC103479452	UN13A	866.972	-0.478	0.001
rna13371|gene=znf746	znf746	1623.600	0.312	0.001
MSTRG.10067.35|gene=MSTRG.10067	KCC2B	7708.760	0.138	0.002
MSTRG.5587.1|gene=MSTRG.5587	NA	387.393	0.325	0.002
rna24078|gene=LOC103472856	SHSA7	1873.094	0.380	0.002
rna4922|gene=LOC103461631	PTPRN	6258.089	0.342	0.002
rna30660|gene=rgs7	rgs7	2386.077	-0.155	0.002
rna46774|gene=abca2	abca2	985.996	0.437	0.002
rna44214|gene=stxbp6	stxbp6	1369.630	0.560	0.002
MSTRG.14206.1|gene=MSTRG.14206	SMBT1	81.425	-8.926	0.002
rna31627|gene=LOC103477235	BTBD3	5069.169	0.085	0.003
MSTRG.22324.1|gene=MSTRG.22324	INT7	201.560	0.525	0.003
rna15437|gene=LOC103467579	IQEC1	262.086	-0.619	0.003
rna30743|gene=LOC103476767	PTPRE	830.200	-0.300	0.003
MSTRG.23588.1|gene=MSTRG.23588	SC5A1	268.214	0.409	0.003
rna18718|gene=sema4c	sema4c	225.613	-0.778	0.003

rna18418|gene=LOC103469455	ZN208	964.520	0.791	0.004
MSTRG.24653.1|gene=MSTRG.24653	ZN665	647.821	-0.301	0.004
rna7261|gene=prc1	prc1	69.432	-0.673	0.004
rna33374|gene=itga8	itga8	177.778	-0.492	0.006
rna14054|gene=LOC103466873	CPNE9	1292.066	-0.186	0.006
rna25858|gene=add1	add1	301.934	-1.200	0.007
rna41611|gene=abracl	abracl	53.205	7.912	0.007
rna8442|gene=tbc1d23	tbc1d23	532.141	0.310	0.009
MSTRG.15056.1|gene=MSTRG.15056	ST32A	275.300	0.492	0.009
rna2302|gene=LOC103472525	NEUL1	298.298	-0.452	0.009
MSTRG.10567.1|gene=MSTRG.10567	EDIL3	901.674	0.304	0.009
rna20215|gene=klf9	klf9	230.772	0.515	0.009
MSTRG.24002.1|gene=MSTRG.24002	KCNA1	5020.815	0.142	0.009
rna43122|gene=npas3	npas3	579.387	0.533	0.010
MSTRG.11745.1|gene=MSTRG.11745	P33MX	3488.184	0.108	0.010
rna28848|gene=gng8	gng8	189.567	-0.795	0.011
MSTRG.14250.1|gene=MSTRG.14250	TMTC3	9253.455	0.190	0.011
rna32670|gene=LOC103477783	PNRC2	6535.494	0.077	0.011
rna37128|gene=LOC103480432	COR1B	96.683	-7.060	0.011
rna836|gene=LOC103465321	PLAK	165.686	-2.434	0.012
MSTRG.22229.1|gene=MSTRG.22229	NA	989.886	0.301	0.012
rna14269|gene=LOC103466973	MYT1	150.362	-8.518	0.012
rna26517|gene=LOC103474357	NA	680.176	0.312	0.013
rna14680|gene=tardbp	tardbp	786.594	-0.256	0.015
rna18417|gene=LOC103469455	ZN208	987.995	-0.992	0.016
rna12053|gene=rpl27a	rpl27a	5598.155	-0.121	0.016
rna3113|gene=fam171b	fam171b	676.632	-0.622	0.017
rna10216|gene=LOC103464678	S12A5	465.767	-0.326	0.017
MSTRG.24178.2|gene=MSTRG.24178	AEBP2	406.228	-0.382	0.018
rna41338|gene=clg20hxorf23	clg20hxorf 23	38.754	-5.680	0.018
rna25124|gene=rpl37	rpl37	3795.051	-0.136	0.018
rna21184|gene=dusp1	dusp1	601.109	0.245	0.019
rna22048|gene=LOC103471326	NA	5028.797	0.222	0.021
MSTRG.7705.1|gene=MSTRG.7705	TADBP	659.341	0.309	0.021
rna15823|gene=LOC103467371	LFG2	6610.378	0.128	0.021
rna42117|gene=foxo3	foxo3	4839.429	0.112	0.021
rna15369|gene=baz2a	baz2a	2468.169	0.271	0.023
rna16768|gene=wbp11	wbp11	637.220	0.247	0.023
rna36928|gene=dgkk	dgkk	330.445	-0.856	0.024
rna8044|gene=uchl5	uchl5	1901.123	-0.123	0.025
rna12667|gene=erc1	erc1	1227.615	0.223	0.026

MSTRG.20354.5|gene=MSTRG.20354	DLG5	269.626	0.856	0.029
rna2024|gene=LOC103470494	GRIA2	2812.527	0.166	0.030
rna10139|gene=pdhb	pdhb	2640.130	-0.139	0.030
MSTRG.8108.1|gene=MSTRG.8108	143BA	11583.530	0.122	0.030
MSTRG.11772.3|gene=MSTRG.11772	RNF6	375.432	0.328	0.030
rna35260|gene=wdr47	wdr47	2309.695	-0.136	0.030
MSTRG.1635.3|gene=MSTRG.1635	NLGNX	1082.247	0.316	0.030
MSTRG.6043.23|gene=MSTRG.6043	NFASC	4875.050	0.157	0.031
rna19015|gene=LOC103469839	SNF5	558.868	-0.187	0.031
rna30678|gene=LOC103476732	ANK3	748.739	0.483	0.031
rna26037|gene=LOC103473957	STXB1	55589.294	0.068	0.031
rna27530|gene=LOC103474859	H33	10943.188	-0.110	0.034
rna8930|gene=LOC103463612	RORB	1423.800	0.184	0.034
rna31916|gene=rps27a	rps27a	10460.725	-0.096	0.035
rna41322|gene=eif2s3	eif2s3	2821.758	-0.139	0.035
rna29177|gene=LOC103475891	AIFM1	41.012	5.526	0.035
MSTRG.25212.1|gene=MSTRG.25212	PLXA3	597.785	0.488	0.037
MSTRG.10784.1|gene=MSTRG.10784	RYR1	25.899	1.375	0.039
rna6858|gene=LOC103462628	CHST8	1822.516	0.123	0.039
rna19888|gene=LOC103470302	NA	327.381	-0.746	0.040
MSTRG.11901.1|gene=MSTRG.11901	NA	1157.565	0.231	0.044
rna28113|gene=actn4	actn4	1503.002	-0.194	0.044
MSTRG.24804.1|gene=MSTRG.24804	ABCA2	1009.735	-0.353	0.045
MSTRG.15875.1|gene=MSTRG.15875	ARHGC	511.656	0.253	0.045
rna1297|gene=fbxo41	fbxo41	2701.441	0.134	0.046
rna11314|gene=LOC103465330	AGAP2	872.129	0.343	0.047
rna25393|gene=cdc37l1	cdc37l1	253.320	-0.270	0.047
rna43115|gene=akap6	akap6	659.763	4.534	0.047
MSTRG.9250.1|gene=MSTRG.9250	MFS6L	901.045	0.193	0.048
rna22955|gene=gatad2b	gatad2b	5809.079	0.156	0.048
MSTRG.1327.2|gene=MSTRG.1327	G137C	1064.576	0.197	0.048
MSTRG.17605.2|gene=MSTRG.17605	CEP72	4481.235	0.158	0.048
MSTRG.9514.1|gene=MSTRG.9514	KPCA	2230.300	0.158	0.049

OPTIC TECTUM

CONTROL LINES (n=187)

Transcript	Gene BLAST	Average expression	Fold Change	P-value (FDR)
rna31156|gene=LOC103476975	ERLN1	44.054	-23.947	< 0.001
rna45754|gene=kcnc2	kcnc2	862.523	0.336	< 0.001
MSTRG.3467.33|gene=MSTRG.3467	MADD	535.023	-0.426	< 0.001
MSTRG.15801.1|gene=MSTRG.15801	FEZ1	709.036	0.450	< 0.001

rna47175|gene=fsd1l	fsd1l	57.585	-6.841	< 0.001
rna8739|gene=LOC103463733	PAIRB	184.831	-8.877	0.001
rna11571|gene=nfasc	nfasc	204.340	-10.374	0.001
rna21659|gene=gsr	gsr	152.518	-0.791	0.001
MSTRG.14289.2|gene=MSTRG.14289	CLPB	860.114	-0.369	0.004
MSTRG.18014.5|gene=MSTRG.18014	SAC1A	583.077	-0.301	0.004
rna38083|gene=eme2	eme2	298.137	-0.519	0.004
MSTRG.5196.2|gene=MSTRG.5196	HEMK1	387.510	-1.001	0.004
rna28725|gene=proser1	proser1	414.492	-0.419	0.004
MSTRG.15730.1|gene=MSTRG.15730	MAP6	1424.326	-0.361	0.004
rna28934|gene=LOC103475715	NPAS2	236.645	-1.310	0.004
MSTRG.917.1|gene=MSTRG.917	MYG	2925.534	0.261	0.004
MSTRG.9656.1|gene=MSTRG.9656	MFS11	246.930	-0.343	0.005
rna12640|gene=sephs1	sephs1	61.114	-8.344	0.005
rna9975|gene=timp4	timp4	4018.568	0.133	0.005
rna21336|gene=sparc	sparc	2184.843	0.289	0.005
rna18695|gene=LOC103469766	XNIF	6926.725	-0.137	0.005
MSTRG.13682.4|gene=MSTRG.13682	DYN1	902.700	-0.436	0.005
rna39368|gene=LOC103481895	PHIPL	1250.523	0.275	0.006
MSTRG.24446.1|gene=MSTRG.24446	NA	129.841	-0.756	0.006
rna20093|gene=LOC103470431	rsmG	56982.959	0.171	0.006
MSTRG.10409.5|gene=MSTRG.10409	HIP1R	229.374	-1.111	0.007
MSTRG.24774.24|gene=MSTRG.24774	E41L3	222.455	-0.670	0.008
rna17061|gene=cacna1i	cacna1i	967.348	-0.420	0.009
rna19619|gene=pde4d	pde4d	389.720	-0.416	0.009
rna23861|gene=clk2	clk2	1467.038	-0.164	0.009
rna20745|gene=LOC103470831	S12A2	1144.392	0.736	0.009
MSTRG.22864.5|gene=MSTRG.22864	MABP1	646.420	-0.385	0.009
rna39712|gene=cacna1g	cacna1g	915.637	-0.347	0.010
MSTRG.10346.1|gene=MSTRG.10346	PERF	664.481	0.237	0.010
MSTRG.24190.1|gene=MSTRG.24190	NYNRI	734.469	-0.722	0.010
MSTRG.13834.2|gene=MSTRG.13834	NMDZ1	4878.659	-0.246	0.010
rna46689|gene=LOC103460147	E41L3	326.656	-0.451	0.010
rna1478|gene=suclg1	suclg1	310.395	0.424	0.010
rna18243|gene=LOC103469344	RHG23	663.109	0.551	0.010
MSTRG.16227.1|gene=MSTRG.16227	RIBA2	133.475	-0.668	0.010
rna11521|gene=LOC103465433	NA	3055.567	-0.322	0.010
rna32056|gene=kiaa0319	kiaa0319	425.023	-0.831	0.011
rna33697|gene=LOC103478428	FEZ2	157.395	-0.446	0.011
rna39112|gene=LOC103481688	ANK3	1151.108	-0.739	0.011
MSTRG.7552.1|gene=MSTRG.7552	GBRD	924.733	-0.409	0.011

rna6818|gene=chd9	chd9	506.594	1.474	0.012
MSTRG.6879.2|gene=MSTRG.6879	AEBP2	755.567	-0.594	0.013
rna8820|gene=bsg	bsg	5022.883	0.117	0.013
rna32896|gene=rbp5	rbp5	700.965	0.314	0.013
rna9941|gene=LOC103464464	NA	6700.276	0.142	0.013
rna4264|gene=s100b	s100b	44245.997	0.203	0.013
rna10292|gene=LOC103464714	SYT2	951.110	-0.683	0.013
rna32087|gene=LOC103477466	PI4KB	1081.824	-0.409	0.013
rna24933|gene=LOC103473353	SMCA2	1265.118	-0.322	0.013
rna8437|gene=gtpbp6	gtpbp6	407.444	-0.307	0.013
MSTRG.20719.13|gene=MSTRG.20719	SHSA6	1502.578	-0.408	0.013
rna43889|gene=LOC103458562	CO052	769.649	-0.511	0.013
rna31795|gene=rims1	rims1	871.867	-0.267	0.013
rna39949|gene=arhgef4	arhgef4	391.940	-0.310	0.013
MSTRG.14169.1|gene=MSTRG.14169	APBP2	489.681	-0.306	0.013
MSTRG.17087.27|gene=MSTRG.17087	SYGP1	2095.222	-0.290	0.013
MSTRG.18474.3|gene=MSTRG.18474	GDE	1577.365	-0.197	0.014
rna24455|gene=ago4	ago4	805.347	0.399	0.014
MSTRG.15686.3|gene=MSTRG.15686	SRRM3	873.289	-0.377	0.014
rna24998|gene=fnbp1	fnbp1	733.773	-0.604	0.015
rna27624|gene=serpine2	serpine2	3522.939	0.124	0.015
rna35717|gene=LOC103479568	PSA7	1970.107	0.113	0.015
rna5025|gene=LOC103460921	PDXK	3641.298	0.200	0.016
MSTRG.24191.1|gene=MSTRG.24191	EFC1	758.821	-0.790	0.016
MSTRG.19248.1|gene=MSTRG.19248	NALP1	250.455	-0.743	0.016
rna14471|gene=LOC103467107	VWA1	291.946	0.459	0.016
rna27096|gene=sik2	sik2	1497.974	-0.340	0.016
rna4809|gene=dapl1	dapl1	529.088	0.361	0.016
rna20030|gene=LOC103470390	LCAP	1033.226	0.321	0.017
rna385|gene=LOC103474313	HEBP2	6881.290	0.112	0.018
rna14585|gene=LOC103467181	GSTM1	383.607	0.332	0.018
MSTRG.1958.1|gene=MSTRG.1958	LRRF2	1052.046	-0.222	0.021
rna16860|gene=pick1	pick1	224.937	-0.839	0.021
MSTRG.9169.3|gene=MSTRG.9169	TTYH3	243.157	0.339	0.021
MSTRG.3724.24|gene=MSTRG.3724	SHAN2	960.854	-0.308	0.022
rna29486|gene=ubb	ubb	1938.270	0.196	0.023
rna35679|gene=LOC103479583	DPP6	226.201	-0.421	0.024
rna1353|gene=pgm2	pgm2	298.878	1.051	0.024
rna4718|gene=ndufs1	ndufs1	704.549	0.317	0.024
rna25185|gene=LOC103473483	SCRB2	1298.625	0.227	0.024
MSTRG.11218.1|gene=MSTRG.11218	CC142	510.828	-0.225	0.024

rna6428|gene=LOC103462451	TRI66	691.559	0.489	0.024
rna39783|gene=atad1	atad1	224.854	-0.367	0.024
rna39263|gene=LOC103481783	ZN235	294.711	-0.637	0.025
rna31608|gene=LOC103477221	CYT	2957.013	0.211	0.027
rna8962|gene=LOC103463594	ZFYV9	2108.108	-0.195	0.027
rna30529|gene=LOC103476656	GDS1	1925.695	0.332	0.027
MSTRG.4676.4|gene=MSTRG.4676	DOT1L	1214.736	-0.288	0.027
rna25497|gene=LOC103473681	NA	2400.413	0.321	0.027
rna33379|gene=LOC103478617	HEY1	707.379	0.232	0.028
rna42980|gene=gja1	gja1	24024.747	0.156	0.029
rna1352|gene=pgm2	pgm2	430.082	-0.541	0.029
MSTRG.4713.3|gene=MSTRG.4713	ZFYV9	3966.457	-0.290	0.029
rna5711|gene=mrpl46	mrpl46	342.311	0.276	0.029
MSTRG.6387.2|gene=MSTRG.6387	AASS	519.728	-0.348	0.029
rna9652|gene=ccdc51	ccdc51	89.289	0.739	0.029
MSTRG.4313.1|gene=MSTRG.4313	LORF2	40.998	1.225	0.029
MSTRG.14276.7|gene=MSTRG.14276	HIP1	502.571	-0.272	0.029
MSTRG.6781.1|gene=MSTRG.6781	NA	176.637	-0.375	0.029
rna34295|gene=aldh2	aldh2	948.974	0.186	0.029
MSTRG.8952.3|gene=MSTRG.8952	CAC1I	2257.156	-0.189	0.030
rna16138|gene=LOC103468055	L1CAM	2054.282	-0.230	0.030
rna17402|gene=rhbdf1	rhbdf1	461.449	-0.505	0.030
rna45|gene=fbxl18	fbxl18	1526.888	-0.201	0.030
MSTRG.21801.2|gene=MSTRG.21801	NA	161.643	-11.290	0.030
MSTRG.11045.2|gene=MSTRG.11045	STAG2	696.471	-0.327	0.030
MSTRG.18008.6|gene=MSTRG.18008	RBM42	981.920	-0.292	0.030
MSTRG.5146.19|gene=MSTRG.5146	CAPS1	2211.292	-0.368	0.031
rna554|gene=LOC103464924	NA	1135.002	0.275	0.031
MSTRG.22927.5|gene=MSTRG.22927	PANK2	199.419	-0.358	0.031
MSTRG.4337.1|gene=MSTRG.4337	NA	163.807	-0.398	0.031
MSTRG.6421.1|gene=MSTRG.6421	DNM1L	284.253	-0.331	0.032
MSTRG.20883.51|gene=MSTRG.20883	CAC1G	2176.919	-0.280	0.032
rna17397|gene=hbz	hbz	2654.700	0.777	0.032
rna9952|gene=pax7	pax7	2715.333	0.252	0.033
MSTRG.12285.7|gene=MSTRG.12285	CSPG5	498.929	-0.448	0.033
MSTRG.4508.3|gene=MSTRG.4508	ANR29	148.773	-0.558	0.034
rna15287|gene=LOC103467653	MYPT2	161.547	-0.529	0.034
MSTRG.21917.2|gene=MSTRG.21917	AIG1	812.919	-0.226	0.034
rna31978|gene=cul9	cul9	815.724	-0.454	0.034
MSTRG.5079.1|gene=MSTRG.5079	CCD66	158.970	-0.655	0.034
rna33477|gene=hdac9	hdac9	1208.808	-0.204	0.034

rna32208|gene=LOC103477566	NA	136.825	-8.874	0.034
MSTRG.23173.1|gene=MSTRG.23173	M3K9	393.992	-0.443	0.034
rna44660|gene=rcan3	rcan3	1297.437	0.176	0.034
MSTRG.22663.6|gene=MSTRG.22663	CNIH3	1399.089	-0.365	0.035
rna2408|gene=LOC103472140	STRN	534.708	-0.374	0.035
MSTRG.12275.1|gene=MSTRG.12275	NA	289.805	-0.465	0.035
MSTRG.11276.1|gene=MSTRG.11276	NA	808.305	0.259	0.035
MSTRG.6086.4|gene=MSTRG.6086	SETD5	646.909	-0.482	0.035
rna20492|gene=ncs1	ncs1	1548.351	0.135	0.035
rna44860|gene=mtpn	mtpn	3939.993	0.096	0.035
rna9563|gene=LOC103464301	SC6A6	3463.148	0.093	0.035
MSTRG.20597.113|gene=MSTRG.20597	ANK3	3978.012	-0.257	0.035
rna15537|gene=LOC103467937	ST1S3	92.866	0.506	0.035
MSTRG.2919.2|gene=MSTRG.2919	KPBB	1982.178	-0.275	0.035
MSTRG.8196.8|gene=MSTRG.8196	CDK17	523.422	-0.384	0.035
MSTRG.10756.1|gene=MSTRG.10756	CACP	73.826	1.141	0.036
MSTRG.21665.4|gene=MSTRG.21665	ASAP1	2262.217	-0.338	0.036
rna9074|gene=LOC103463535	SAFB2	486.492	0.877	0.036
MSTRG.6893.1|gene=MSTRG.6893	KCND2	751.620	-0.249	0.037
rna37092|gene=gabarap	gabarap	9554.826	0.122	0.037
MSTRG.21609.4|gene=MSTRG.21609	MYPT1	538.478	-1.474	0.037
MSTRG.17335.7|gene=MSTRG.17335	KPCA	1293.373	-0.233	0.037
MSTRG.14062.6|gene=MSTRG.14062	NCAM1	671.251	-0.400	0.038
rna48393|gene=LOC103461400	PPAL	460.373	0.240	0.039
rna43205|gene=ckb	ckb	568.475	0.225	0.039
rna46432|gene=ddx46	ddx46	2735.921	-0.244	0.039
rna10428|gene=LOC103464779	HES5	738.767	0.267	0.039
rna13605|gene=LOC103466450	RFX7	1889.887	0.384	0.039
rna1698|gene=LOC103468483	ELOF1	256.797	0.280	0.039
rna37063|gene=bcl6b	bcl6b	201.983	0.501	0.039
rna36246|gene=smg7	smg7	2826.022	-0.149	0.039
MSTRG.8764.2|gene=MSTRG.8764	DDX5	9210.083	-0.235	0.039
rna20266|gene=LOC103470511	DNM1L	1990.942	-0.169	0.039
MSTRG.8986.28|gene=MSTRG.8986	CAC1A	5587.949	-0.207	0.039
MSTRG.12290.2|gene=MSTRG.12290	CTNB1	2910.930	-0.206	0.039
rna564|gene=gde1	gde1	785.583	0.260	0.041
MSTRG.16875.3|gene=MSTRG.16875	S22AF	106.886	-0.823	0.041
MSTRG.11867.7|gene=MSTRG.11867	GRIA2	7579.132	-0.254	0.041
rna28723|gene=thsd1	thsd1	124.740	-0.707	0.042
rna32852|gene=casp2	casp2	194.575	0.487	0.042
MSTRG.19932.2|gene=MSTRG.19932	ERF3A	147.054	-0.532	0.042

MSTRG.14143.1|gene=MSTRG.14143	WSB1	2604.092	-0.329	0.043
MSTRG.1695.1|gene=MSTRG.1695	WBP4	118.747	-0.790	0.043
rna12783|gene=cib1	cib1	172.780	0.314	0.045
rna39046|gene=LOC103481669	RHG27	219.443	0.377	0.045
rna4111|gene=etv5	etv5	177.296	0.799	0.045
rna45179|gene=LOC103459268	CAC1C	1903.168	-0.854	0.046
rna43863|gene=zbtb42	zbtb42	95.269	-10.618	0.047
rna43182|gene=tyro3	tyro3	1426.998	0.392	0.047
rna26654|gene=snx19	snx19	2368.539	-0.171	0.047
rna16854|gene=LOC103468448	BORG5	259.020	0.308	0.048
MSTRG.8096.18|gene=MSTRG.8096	WNK2	3178.569	-0.321	0.049
MSTRG.17167.1|gene=MSTRG.17167	ARL4A	107.263	-9.441	0.049
MSTRG.10976.1|gene=MSTRG.10976	NA	339.637	-0.260	0.049
rna18354|gene=LOC103469494	EIF1	480.176	-0.425	0.049
rna12804|gene=slc1a2	slc1a2	38882.612	0.102	0.049
MSTRG.5171.1|gene=MSTRG.5171	RBM5	1463.233	-0.303	0.049
MSTRG.18686.1|gene=MSTRG.18686	RHG12	753.715	-0.292	0.049
MSTRG.24504.1|gene=MSTRG.24504	RAE1	743.520	-0.294	0.049
MSTRG.3297.5|gene=MSTRG.3297	DGLA	461.380	-0.328	0.049

POLARIZATION-SELECTED LINES (n=21)

Transcript	Gene BLAST	Average expression	Fold Change	P-value (FDR)
MSTRG.21801.2|gene=MSTRG.21801	NA	102.640	24.315	< 0.001
rna20241|gene=entpd4	entpd4	13.232	21.501	< 0.001
rna5604|gene=cdh8	cdh8	1553.154	0.643	< 0.001
rna24376|gene=scn1b	scn1b	92.098	19.494	< 0.001
rna22214|gene=egr1	egr1	39.104	18.875	< 0.001
rna22213|gene=egr1	egr1	39.104	18.875	< 0.001
MSTRG.2915.2|gene=MSTRG.2915	CADH8	292.560	-7.961	< 0.001
rna15769|gene=LOC103467406	DSLE	156.024	-0.735	0.002
MSTRG.22698.1|gene=MSTRG.22698	LAMA4	269.181	-8.539	0.004
rna29823|gene=wwc1	wwc1	283.825	-9.054	0.008
rna1269|gene=gucy1a3	gucy1a3	206.417	-0.773	0.010
rna28422|gene=LOC103475388	CTR3	1020.217	0.605	0.015
rna35022|gene=LOC103479254	KC1G2	174.252	-1.351	0.020
rna16375|gene=LOC103468166	MPP2	287.359	0.758	0.020
rna29120|gene=LOC103475822	SPTB2	245.255	-8.404	0.032
rna12010|gene=slc6a15	slc6a15	4110.120	0.134	0.034
rna23231|gene=ahdc1	ahdc1	875.602	-0.421	0.034
rna35400|gene=LOC103479433	P66B	129.015	-0.435	0.034

rna39880|gene=tmem94	tmem94	138.800	8.215	0.040
rna7779|gene=LOC103463266	RAB3C	271.999	-0.557	0.040
rna41072|gene=LOC103456849	ESCO1	357.308	11.958	0.040

TELENCEPHALON

CONTROL LINES (n=29)

Transcript	Gene BLAST	Average expression	Fold Change	P-value (FDR)
rna16008|gene=zmynd12	zmynd12	131.990	3< 0.001	< 0.001
rna45464|gene=slc37a3	slc37a3	38.721	-22.872	< 0.001
rna14004|gene=LOC103466852	CNGB1	40.820	22.427	< 0.001
rna36260|gene=LOC103480019	PTPR2	26.631	-22.537	< 0.001
MSTRG.16482.1|gene=MSTRG.16482	MORN4	15.553	-22.396	< 0.001
rna9271|gene=ncln	ncln	35.830	8.337	< 0.001
rna1296|gene=fbxo41	fbxo41	364.240	-0.593	< 0.001
rna10792|gene=plxna1	plxna1	299.891	1.086	0.001
rna34474|gene=LOC103478864	HHATL	249.894	0.719	0.003
MSTRG.16496.1|gene=MSTRG.16496	LDB1	437.207	-0.439	0.006
rna8510|gene=cc2d2a	cc2d2a	231.497	-0.488	0.007
rna45291|gene=grip1	grip1	1272.284	0.304	0.009
rna564|gene=gde1	gde1	945.463	0.225	0.009
rna752|gene=tsc2	tsc2	572.522	-0.466	0.014
rna33839|gene=LOC103478362	COBL	316.960	0.373	0.015
rna38088|gene=LOC103481105	VATL	241.247	-0.797	0.015
MSTRG.10409.5|gene=MSTRG.10409	HIP1R	136.542	-1.063	0.016
MSTRG.20355.9|gene=MSTRG.20355	KCMA1	1362.688	-0.259	0.017
MSTRG.21801.2|gene=MSTRG.21801	NA	148.516	-8.653	0.018
rna13396|gene=islr2	islr2	11284.671	0.081	0.020
MSTRG.22908.2|gene=MSTRG.22908	ATRN	287.114	-0.510	0.020
MSTRG.22993.1|gene=MSTRG.22993	SSBP3	1070.362	0.254	0.020
rna2770|gene=map4k4	map4k4	1829.434	0.204	0.020
MSTRG.205.3|gene=MSTRG.205	IFT27	76.967	2.301	0.021
rna45311|gene=ptprb	ptprb	221.601	-10.953	0.027
rna28655|gene=LOC103475549	TRIM3	234.203	0.738	0.027
rna26416|gene=LOC103474211	ROBO2	1300.445	-0.301	0.044
MSTRG.23960.3|gene=MSTRG.23960	RB6I2	643.993	-0.251	0.046
rna16561|gene=LOC103468305	FRIM	3458.490	-0.117	0.049

POLARIZATION-SELECTED LINES (n=158)

Transcript	Gene BLAST	Average expression	Fold Change	P-value (FDR)
rna18687|gene=tia1	tia1	401.711	0.422	< 0.001

MSTRG.23945.1|gene=MSTRG.23945	CAC1C	1299.068	0.768	< 0.001
rna31577|gene=LOC103477205	TACC2	2488.829	0.178	< 0.001
MSTRG.21916.1|gene=MSTRG.21916	ZEP2	6873.343	0.187	< 0.001
rna9693|gene=abhd14a	abhd14a	1497.664	-0.168	0.001
MSTRG.18769.2|gene=MSTRG.18769	KIF1A	7524.576	0.217	0.001
rna20419|gene=rictor	rictor	558.369	0.353	0.001
rna42366|gene=LOC103457350	MAX	355.081	0.441	0.001
MSTRG.7449.1|gene=MSTRG.7449	AT2B3	4327.516	0.172	0.001
MSTRG.11745.1|gene=MSTRG.11745	P33MX	3551.885	-0.160	0.002
rna15494|gene=LOC103467551	SLMAP	1057.196	-0.220	0.002
rna10910|gene=LOC103464892	SORT1	2240.260	0.317	0.002
rna46476|gene=ppm1e	ppm1e	1440.302	0.315	0.003
rna39999|gene=LOC103482255	PFKAP	882.019	-1.088	0.003
rna39992|gene=LOC103482255	PFKAP	1367.223	0.413	0.003
rna9416|gene=LOC103464216	LRC24	586.747	0.403	0.003
MSTRG.9631.2|gene=MSTRG.9631	ADA11	7045.660	0.126	0.003
rna42290|gene=gphn	gphn	1236.892	0.239	0.006
rna139|gene=pcm1	pcm1	415.579	0.513	0.006
MSTRG.26314.1|gene=MSTRG.26314	LIMS1	248.767	-0.421	0.006
MSTRG.9224.4|gene=MSTRG.9224	GFAP	2397.131	-0.223	0.008
rna21649|gene=atp8a1	atp8a1	223.183	7.673	0.008
rna30694|gene=pgbd5	pgbd5	1308.949	0.466	0.008
MSTRG.14919.1|gene=MSTRG.14919	KLC1	317.188	-0.494	0.009
rna44843|gene=magi2	magi2	333.647	-6.397	0.009
rna4264|gene=s100b	s100b	8575.720	-0.294	0.011
rna8601|gene=LOC103463821	REM2	1266.508	0.215	0.011
MSTRG.5319.1|gene=MSTRG.5319	MTG8R	289.321	0.659	0.011
MSTRG.8385.2|gene=MSTRG.8385	CHDH	613.368	-0.337	0.011
rna35142|gene=LOC103479283	PTPRS	3184.232	0.228	0.011
MSTRG.25460.5|gene=MSTRG.25460	DPYL2	1490.984	-0.159	0.011
rna10217|gene=LOC103464678	S12A5	292.625	0.760	0.011
rna3294|gene=LOC103477739	T22D1	942.977	0.314	0.013
MSTRG.16563.2|gene=MSTRG.16563	KHDR2	596.821	-0.263	0.014
MSTRG.904.5|gene=MSTRG.904	SMCA4	825.667	0.940	0.014
rna7625|gene=LOC103463153	KCC4	813.768	0.306	0.015
rna23887|gene=LOC103472687	RS27	8841.463	-0.132	0.015
rna40585|gene=atp6v1a	atp6v1a	3486.538	0.279	0.015
rna28612|gene=LOC103475568	NBEA	1738.565	0.230	0.015
rna30893|gene=eif4ebp2	eif4ebp2	4206.554	-0.129	0.015
rna22733|gene=LOC103472023	SIA4A	329.212	-0.484	0.016
rna22398|gene=gria2	gria2	1501.733	0.394	0.018

MSTRG.13007.1|gene=MSTRG.13007	DEN1A	523.827	-0.258	0.018
MSTRG.604.1|gene=MSTRG.604	PCD10	853.615	-0.214	0.020
rna20270|gene=LOC103470513	BMP1	704.668	0.216	0.020
MSTRG.8142.2|gene=MSTRG.8142	GRM7	2211.218	0.219	0.020
rna26271|gene=apbb1	apbb1	616.593	-0.529	0.021
rna7160|gene=LOC103463007	SHAN1	1246.491	0.303	0.021
rna38233|gene=jph3	jph3	1481.033	0.154	0.021
rna9994|gene=atp2b2	atp2b2	2968.692	0.160	0.021
MSTRG.7852.1|gene=MSTRG.7852	TM269	642.667	-0.332	0.021
rna41217|gene=u2surp	u2surp	260.347	0.496	0.021
rna16889|gene=smcr8	smcr8	505.442	0.242	0.021
rna20696|gene=LOC103470760	ODO1	3680.969	0.128	0.022
rna6714|gene=madd	madd	581.771	0.781	0.024
rna28842|gene=LOC103475446	PP2BA	1272.204	0.483	0.025
rna18171|gene=crebbp	crebbp	477.352	0.698	0.025
MSTRG.24874.5|gene=MSTRG.24874	LHPL5	1008.636	-0.240	0.025
rna7423|gene=LOC103463224	DDAH1	614.691	-0.294	0.025
rna38076|gene=LOC103481097	MPRIP	1004.516	0.694	0.026
rna15448|gene=bicd2	bicd2	802.184	0.257	0.026
rna21869|gene=mid2	mid2	858.588	-0.264	0.028
rna1133|gene=LOC103463598	S14L1	402.061	0.463	0.028
rna18056|gene=tom1l2	tom1l2	1234.140	-0.201	0.028
rna26413|gene=LOC103474211	ROBO2	957.985	0.457	0.029
rna3034|gene=tmem47	tmem47	663.543	-0.408	0.029
MSTRG.26426.1|gene=MSTRG.26426	AT2B3	509.885	0.223	0.029
rna14675|gene=kif1b	kif1b	6926.073	0.221	0.029
rna28593|gene=pds5b	pds5b	4058.133	0.116	0.029
rna6897|gene=LOC103462737	RGRF1	3742.185	0.223	0.029
rna12565|gene=cacna2d1	cacna2d1	307.267	0.642	0.029
rna25879|gene=LOC103473860	SHC3	917.756	0.542	0.030
MSTRG.20400.1|gene=MSTRG.20400	NETR	2933.854	-1.107	0.030
rna10085|gene=r3hdm2	r3hdm2	882.792	0.293	0.030
rna2232|gene=sh3gl2	sh3gl2	1153.102	0.336	0.030
rna3256|gene=LOC103477319	NA	369.700	-0.551	0.030
rna7150|gene=cttn	cttn	1432.342	0.154	0.030
MSTRG.23582.7|gene=MSTRG.23582	UNC79	2182.014	-0.299	0.031
rna28372|gene=LOC103475354	SEP8A	1479.812	0.221	0.031
rna40511|gene=stau2	stau2	3751.772	0.101	0.031
rna18955|gene=rasgef1b	rasgef1b	314.812	0.469	0.031
rna40847|gene=cnksr2	cnksr2	2782.045	0.236	0.031
rna32682|gene=LOC103477774	NA	5094.213	-0.150	0.031

rna11361|gene=LOC103465345	LRP1	1464.591	0.319	0.031
rna12226|gene=ptprz1	ptprz1	939.568	-0.197	0.032
MSTRG.24892.1|gene=MSTRG.24892	JUN	1541.880	0.159	0.032
rna48421|gene=LOC103461426	PACA	6547.169	0.114	0.032
MSTRG.1426.7|gene=MSTRG.1426	MCF2L	1541.030	0.216	0.032
rna19642|gene=smarcad1	smarcad1	374.477	0.279	0.033
rna45583|gene=LOC103459444	PRIC1	953.356	0.188	0.033
rna9152|gene=ttll7	ttll7	351.793	0.344	0.033
rna9558|gene=rims4	rims4	2057.525	0.123	0.033
rna26824|gene=baz1b	baz1b	250.301	0.419	0.033
rna27920|gene=LOC103475082	CUL5	478.350	0.422	0.034
rna30614|gene=LOC103476702	PKHA1	1179.339	-0.144	0.034
MSTRG.10701.2|gene=MSTRG.10701	DPYL2	293.373	-0.561	0.035
rna22264|gene=LOC103471194	SFXN1	1299.021	-0.161	0.035
rna45277|gene=cand1	cand1	2417.395	0.115	0.035
rna44024|gene=srrm1	srrm1	719.011	0.472	0.035
rna14390|gene=gnb1	gnb1	2102.920	0.133	0.035
rna19155|gene=camsap1	camsap1	366.586	-0.849	0.035
rna11115|gene=LOC103465107	BSN	3020.560	0.361	0.038
rna13371|gene=znf746	znf746	1535.108	0.216	0.038
rna39374|gene=LOC103481898	KGP1	3810.363	0.112	0.038
rna45925|gene=LOC103459684	CDK17	1635.928	0.146	0.038
rna22088|gene=LOC103471305	SSH	1433.568	0.179	0.038
rna9540|gene=LOC103464285	AT2B4	966.284	0.325	0.038
rna7314|gene=LOC103462926	GSLG1	2352.158	0.169	0.039
rna25580|gene=dgcr2	dgcr2	862.995	0.353	0.039
rna19014|gene=ralgds	ralgds	289.546	0.335	0.039
rna46771|gene=rad17	rad17	289.675	-0.260	0.040
rna21459|gene=LOC103471609	ZIC3	848.137	0.211	0.041
rna33285|gene=LOC103478674	T200B	768.169	0.236	0.041
rna29284|gene=klc2	klc2	1027.397	0.157	0.041
rna21130|gene=LOC103471057	BRD8	579.688	0.296	0.041
rna47838|gene=LOC103460885	ATRN1	749.916	0.186	0.042
rna43237|gene=jade1	jade1	486.768	0.251	0.042
rna39189|gene=srcap	srcap	3117.647	0.130	0.043
rna5118|gene=LOC103460319	TM163	2283.059	0.153	0.044
rna40679|gene=arfgef1	arfgef1	1093.914	0.292	0.044
rna36138|gene=LOC108167071	MIB1	579.766	-0.245	0.044
rna33009|gene=gsk3a	gsk3a	390.156	0.245	0.045
rna33558|gene=kmt2b	kmt2b	1100.465	0.334	0.045
rna7409|gene=LOC103463058	LRP8	248.208	0.326	0.045

rna36173|gene=dnajb4	dnajb4	1321.044	0.178	0.045
rna2990|gene=LOC103474271	SCN2A	797.677	0.430	0.045
rna12431|gene=LOC103465918	S6A13	261.539	-0.709	0.045
rna39333|gene=LOC103481878	GRID1	688.931	0.397	0.045
MSTRG.11445.1|gene=MSTRG.11445	TSN7	913.167	-0.359	0.045
rna4111|gene=etv5	etv5	603.824	0.555	0.045
rna14581|gene=LOC103467180	AMPD2	4309.644	0.105	0.045
rna17113|gene=sox8	sox8	1165.511	0.165	0.045
rna6646|gene=LOC103462596	TLN2	1334.136	0.289	0.045
rna1884|gene=acox1	acox1	922.029	-0.142	0.045
MSTRG.11781.2|gene=MSTRG.11781	ATRX	5501.649	0.144	0.045
rna41153|gene=ptprn2	ptprn2	11939.245	0.175	0.045
rna4922|gene=LOC103461631	PTPRN	8049.927	0.206	0.045
rna2054|gene=LOC103470599	TEN3	754.778	-0.460	0.045
rna37531|gene=LOC103480722	PP2BA	3022.883	0.160	0.045
MSTRG.2884.1|gene=MSTRG.2884	CD81	3884.961	-0.100	0.045
rna10431|gene=LOC103464782	HES5	297.115	-0.279	0.045
rna35326|gene=LOC103479394	SCG2	38239.656	0.110	0.045
rna19050|gene=LOC103469848	KCC2B	398.238	0.440	0.046
MSTRG.9859.3|gene=MSTRG.9859	TIA1	386.806	-0.289	0.046
rna2730|gene=app	app	1253.390	0.129	0.046
rna43935|gene=kbtbd11	kbtbd11	3946.524	0.180	0.048
rna22259|gene=gabrb2	gabrb2	2528.544	0.149	0.048
rna28126|gene=hnrnpl	hnrnpl	1361.666	0.189	0.048
rna4001|gene=sumo1	sumo1	1338.464	-0.129	0.049
rna19299|gene=slc20a2	slc20a2	629.265	-0.668	0.049
MSTRG.10579.6|gene=MSTRG.10579	MEF2C	2175.170	-0.305	0.049
rna32394|gene=cspg5	cspg5	7576.875	0.148	0.049
rna12473|gene=syt12	syt12	4226.735	0.156	0.049
rna11522|gene=LOC103465434	CDK18	613.584	0.944	0.049
MSTRG.21665.4|gene=MSTRG.21665	ASAP1	1872.921	-0.200	0.049
MSTRG.8757.1|gene=MSTRG.8757	UBP43	232.268	-0.358	0.049
rna43050|gene=ndufaf4	ndufaf4	394.548	-0.229	0.049
rna28590|gene=fry	fry	756.197	0.748	0.049

Table S7. Summary of differentially co-expressed gene pair results

	OPTIC TECTUM			TELENCEPHALON			MIDBRAIN		
	Control line		Schooling line	Control line		Schooling line	Control line		Schooling line
DC gene pairs	29		65	228		35	67		60
DC unique genes	27		61	161		36	61		55
DC genes in common (between CT and SCH lines)		NA			*ELAV2*			*MKL2* *IGF1R* *TRIM39* *RPL19* *POLR1B*	
Genes that are DC in more than one comparison	*TNIKB* *GBA2*		*ELAVL2* *MSTRG.8823* *RANBP3*	*NTMT1* *RL19* *TRI39*	*LOC103468557* *ARL4A* *FACR1*	*GRIA2* *GOLGA5* *PARVA*	*KCAB1* *IGF1R* *PFKFB2*	*LBH* *MKL2* *PTGR2*	*SURF2* *POLR1B*

Table S8. List of differentially coexpressed genes identified by BFDCA by brain region and line with number of differentially correlated connections (DC gene pairs) for each gene. Genes highlighted in green are DC in more than one comparison, and those highlighted in blue are also Diferentially Expressed.

OPTIC TECTUM				TELENCEPHALON				MIDBRAIN			
Control lines		Polarization-selected lines		Control lines		Polarization-selected lines		Control lines		Polarization-selected lines	
ADPRH	4	*TLK2*	7	mea1	13	*slc17a6*	4	*MAGI1*	5	*LOC103461370*	6
CGAS	4	*AP1B1*	6	KAT7	7	*tenm2*	4	*csk*	4	*PLAK*	6
ABCA1	3	*rtn4*	5	ndufs4	7	*tm2d1*	4	*IGF1R*	4	*mmp16*	5
CENPS	3	*ZN226*	5	apmap	6	*clmn*	3	*PPARA*	4	*CNTNAP5*	4
MSTRG.7285	3	*PARVA*	4	arhgef4	6	*FACR1*	3	*CL17A*	3	*CPSF3*	4
LOC103468557	3	*pfkfb2*	4	NIPLB	6	*KHDR2*	3	*F264*	3	*pou4f1*	4
NRIP3	3	*SHC3*	4	pkn1	6	*LRC24*	3	*hmga2*	3	*abracl*	3
pcdh19	3	*tom1l2*	4	rpl17	6	*PTPRS*	3	*LOC103468544*	3	*CCDC127*	3
STIMA	3	*ASIC1*	3	TRIM69	6	*SUIS_HUMAN*	2	*LOC103468557*	3	*DNAI2*	3
SURF2	3	*CCDC51*	3	ARIP4	5	*grm1*	2	*MILK1*	3	*HES6*	3
BC11B	2	*CDH6*	3	cndp2	5	*LBH*	2	*polr1b*	3	*LOC103461363*	3
DLGAP2	2	*ERLN1*	3	ldhb	5	*MACF1*	2	*RASSF5*	3	*NUP205*	3
ENAH	2	*GRID2*	3	LRRT4	5	*MSTRG.5809*	2	*rbm33*	3	*otp*	3
GBA2	2	*KCC2B*	3	rps8	5	*mthfs*	2	*rbm45*	3	*pdgfrb*	3
KCNH7	2	*KLHL17*	3	smarca4	5	*nrbp2*	2	*S4A7*	3	*ABRAL*	2
mapk6	2	*mphosph8*	3	tbc1d13	5	*PNN*	2	*ABCD1*	2	*ARI1B*	2
rbms1	2	*ntmt1*	3	AT8B2	4	*RAB3A*	2	*dctn1*	2	*ARL4A*	2
slc29a2	2	*r3hdm2*	3	CYTB	4	*rheb*	2	*dnal1*	2	*ARL6*	2
tmbim6	2	*TSPAN8*	3	dvl1	4	*rpl7*	2	*DNPH1*	2	*CALB2*	2
ADHX	1	*cdk14*	2	gatad2b	4	*RXRBA*	2	*DOCK9*	2	*clg20hxorf23*	2
ARL4A	1	*CROL2*	2	htr7	4	*TLK1B*	2	*eaf1*	2	*CORO2A*	2
clns1a	1	*EAF1*	2	ier2	4	*TM205*	2	*GRIA2*	2	*ctps1*	2
SCD5	1	*FOXP2*	2	itgb5	4	*ABLIM2*	1	*grip1*	2	*gapdhs*	2
sephs1	1	*HAPLN1*	2	jarid2	4	*CSK21*	1	*HMCS1*	2	*golga5*	2
slc39a7	1	*HDGFL3*	2	NUCB2	4	*elavl2*	1	*IPKA*	2	*GTDC1*	2
tnikb	1	*HOME3*	2	PCDHGA2	4	*LYSC*	1	*itpkb*	2	*IGF1R*	2
ZN235	1	*LOC103459553*	2	rabgap1	4	*MAPK9*	1	*KCAB1*	2	*KCNJ9*	2
		lrrtm4l1	2	rpl14	4	*MSTRG.8823*	1	*LBH*	2	*LOC103460906*	2
		NPAS2	2	SATT	4	*ntmt1*	1	*LOC103476863*	2	*mkl2*	2
		NRP1A	2	scai	4	*PP2BA*	1	*LORF2*	2	*nrg2*	2
		NSG2	2	SCN5A	4	*ppp1r13l*	1	*LRRF2*	2	*NTM1B*	2
		PHF1	2	tbl1xr1	4	*PTGR2*	1	*map6d1*	2	*polr1b*	2
		ptprd	2	tmem145	4	*thumpd2*	1	*mkl2*	2	*ranbp3*	2
		st3gal4	2	tnrc6b	4	*TRAPPC1*	1	*nfe2l1*	2	*RL19*	2
		ube2v2	2	ZN865	4	*TSP1L*	1	*nlgn2*	2	*rpl37a*	2
		ALDH18A1	1	aars2	3	*zmynd12*	1	*NPRL3*	2	*SEPTIN12*	2
		APBA1	1	ache	3			*pfkp*	2	*SFMBT1*	2
		cfap99	1	AGRN	3			*PKIA*	2	*sin3a*	2
		CHD8	1	AMT	3			*PRC1*	2	*TLE4*	2
		cpeb4	1	ANK2	3			*prickle2*	2	*TRI39*	2
		EIF4G3	1	ASAP2	3			*rexo1*	2	*abcg4*	1
		GLSK	1	ATX1	3			*rpl15*	2	*cdc42bpb*	1
		GLUL	1	bdnf	3			*SURF2*	2	*COR1A*	1
		GNG7	1	bicd2	3			*thap5*	2	*FACR1*	1
		heatr5b	1	cacna1a	3			*tnikb*	2	*gdpd2*	1
		ITPR1	1	cdipt	3			*TRI39*	2	*IL1A*	1
		kiaa2026	1	COX18	3			*TSPAN7*	2	*MRPL55*	1
		LOC103464370	1	cpsf6	3			*TUFT1*	2	*MSTRG.24943*	1
		LOC103477566	1	csgalnact2	3			*ZFP2*	2	*ncoa4*	1
		NCAM2	1	DDX42	3			*atp2b2*	1	*orc4*	1
		NR1AA	1	dnmbp	3			*CNGB1*	1	*PARVA*	1
		nt5e	1	dtnb	3			*DCNL1*	1	*pfkfb2*	1
		orc5	1	fau	3			*elavl4*	1	*pla1a*	1
		ptprm	1	FLRT1	3			*GBA2*	1	*PTGR2*	1
		ranbp3	1	GRIA2	3			*grk4*	1	*RHG22*	1
		rpl8	1	HACE1	3			*MSTRG.8823*	1		
		RTase	1	HDAC4	3			*MYO1E*	1		
		SRSF5	1	HS6ST2	3			*ran*	1		
		TSC22D2	1	hypk	3			*RL19*	1		

vps45	1	KCD15	3	*rundc3a*	1
XPC	1	MAGI2	3	*SCHIP1*	1
		nell1	3	*shroom2*	1
		NTRK2	3	*SYPH*	1
		NYNRIN	3	*tanc2*	1
		P85B	3	*znf512*	1
		PMM2	3		
		psmc3	3		
		RAC3	3		
		RADIL	3		
		RL19	3		
		rpl10	3		
		rpl11	3		
		rpl30	3		
		rpl7a	3		
		rps3a	3		
		SCAM5	3		
		slc1a1	3		
		SPAG9	3		
		sptbn4	3		
		SYGP1	3		
		TACC1	3		
		tmem19	3		
		TRPC4AP	3		
		UBA3	3		
		USP9X	3		
		wdr18	3		
		zcchc6	3		
		zyg11b	3		
		AN13C	2		
		AN32E	2		
		atg3	2		
		B3GL2	2		
		CHD5	2		
		csmd2	2		
		EF1A1	2		
		EVL	2		
		GNAQ	2		
		golga5	2		
		GRM5	2		
		KCNH8	2		
		kdm4a	2		
		LAP2	2		
		LOC103460099	2		
		LOC103467299	2		
		lrrc75a	2		
		LRRT2	2		
		marcksl1	2		
		ncln	2		
		NLGN1	2		
		NOVA1	2		
		NTL9	2		
		PAK1IP1	2		
		pfdn6	2		
		PHLPP2	2		
		ppargc1b	2		
		PPP2R5C	2		
		prkca	2		
		RL12	2		
		rpl31	2		
		rplp0	2		
		S4A11	2		
		SCN8A	2		
		SI1L1	2		
		slc1a2	2		
		slc6a11	2		
		spen	2		
		SPTBN4	2		
		TESK2	2		
		TMTC3	2		

	TRPM6	2	
	uba52	2	
	UBC2	2	
	UNC13A	2	
	XPO1	2	
	ZN665	2	
	ZNF521	2	
	znf839	2	
	AHNK	1	
	BMTI6	1	
	BSN	1	
	ctsa	1	
	elavl2	1	
	GP155	1	
	GPR37	1	
	hspd1	1	
	KCAB1	1	
	LOC103460110	1	
	LOC103460981	1	
	lrig1	1	
	lrrc8a	1	
	MYT1	1	
	ndufa11	1	
	PLXA2	1	
	ptprb	1	
	radil	1	
	SHSA6	1	
	snrk	1	
	snrpa	1	
	SNTG1	1	
	tnikb	1	
	TRI46	1	

Table S9. Summary of genes found to be differentially co-expressed and differentially expressed in our transcriptomic analyses

COMMON DE & DC GENES					
OPTIC TECTUM		TELENCEPHALON		MIDBRAIN	
CONTROL	POLARIZATION	CONTROL	POLARIZATION	CONTROL	POLARIZATION
3	0	2	4	3	0
SEPHS1 *ZN235* *ARL4A*		*NCLN* *PTPRB*	*LRRC24* *PTPRS* *KHDR2* *PP2BA*	*IGF1R* *ITPKB* *TANC2*	

Table S10. KEGG pathways of interest associated in BFDCA differentially coexpressed gene pairs. Only showing pathways and select genes relevant to neural processes associated with behaviour that presented a FDR corrected p-value <0.05

OPTIC TECTUM	
Polarization-selection lines	
Pathways of interest (enriched)	Genes of interest
Glutamatergic synapse	GLUL,GNG7,ITPR1
Long-term depression	GRID2,ITPR1
Dopaminergic synapse	GNG7,ITPR1
GABAergic synapse	GLUL,GNG7
Cholinergic synapse	GNG7,ITPR1
Serotonergic synapse	GNG7,ITPR1
Control lines	
Signal Transduction of S1P Receptor	MAPK6
TELENCEPHALON	
Polarization-selection lines	
Pathways of interest (enriched)	Genes of interest
FoxO signaling pathway	GRM1,MAPK9
Phospholipase D signaling pathway	GRM1,RHEB
Synaptic vesicle cycle	SLC17A6,RAB3A
Glutamatergic synapse	SLC17A6,GRM1
JNK (c-Jun kinases) phosphorylation and activation	MAPK9
mediated by activated human TAK1	RAB3A
GABA synthesis, release, reuptake and degradation	RAB3A
Glutamate Neurotransmitter Release Cycle	
Control lines	
Glutamatergic synapse	CACNA1A,GRIA2,SLC1A1,GNAQ,GRM5,PRKCA,SLC1A2
Synaptic vesicle cycle	CACNA1A,SLC1A1,SLC1A2,SLC6A11,UNC13A
PI3K-Akt signaling pathway	PKN1,ITGB5,BDNF,MAGI2,NTRK2,PHLPP2,PPP2R5C,PRKCA
Long-term depression	CACNA1A,GRIA2,GNAQ,PRKCA
Dopaminergic synapse	CACNA1A,GRIA2,GNAQ,PPP2R5C,PRKCA
Long-term potentiation	GRIA2,GNAQ,GRM5,PRKCA
Cholinergic synapse	ACHE,CACNA1A,GNAQ,PRKCA
Ras signaling pathway	HTR7,BDNF,NTRK2,RAC3,PRKCA
MIDBRAIN	
Polarization-selection lines	
EGFR tyrosine kinase inhibitor resistance	PDGFRB,IGF1R,NRG2
Ras signaling pathway	PDGFRB,IGF1R,PLA1A
MAPK signaling pathway	PDGFRB,IGF1R,IL1A
AMPK signaling pathway	IGF1R,PFKFB2
Control lines	
cAMP signaling pathway	PPARA,GRIA2,ATP2B2,CNGB1
Rap1 signaling pathway	MAGI1,IGF1R,RASSF5
Long-term depression	IGF1R,GRIA2
AMPK signaling pathway	IGF1R,PFKP
cGMP-PKG signaling pathway	ATP2B2,CNGB1

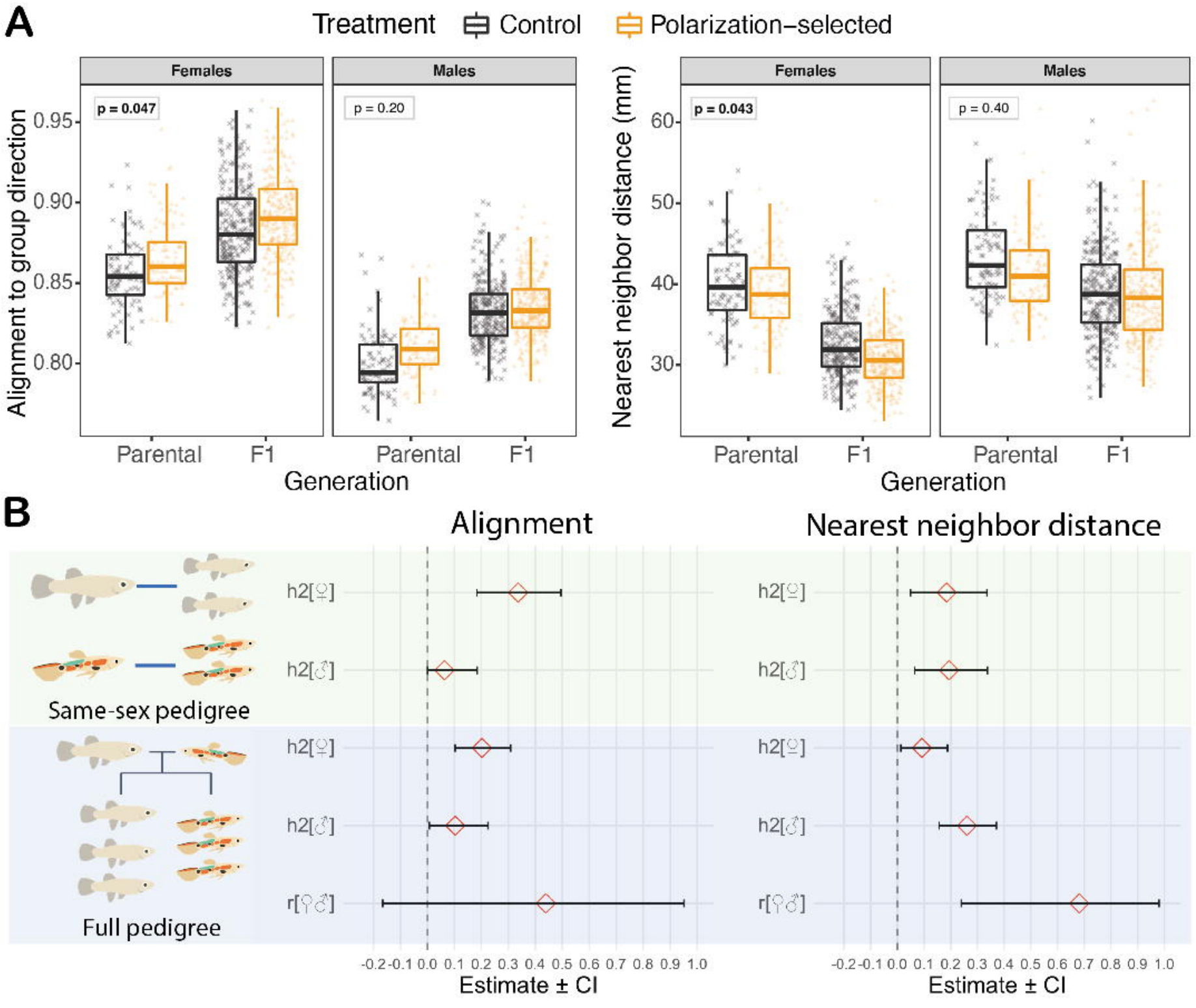
Glutamate binding, activation of AMPA receptors and synaptic plasticity	GRIA2,GRIP1,TSPAN7

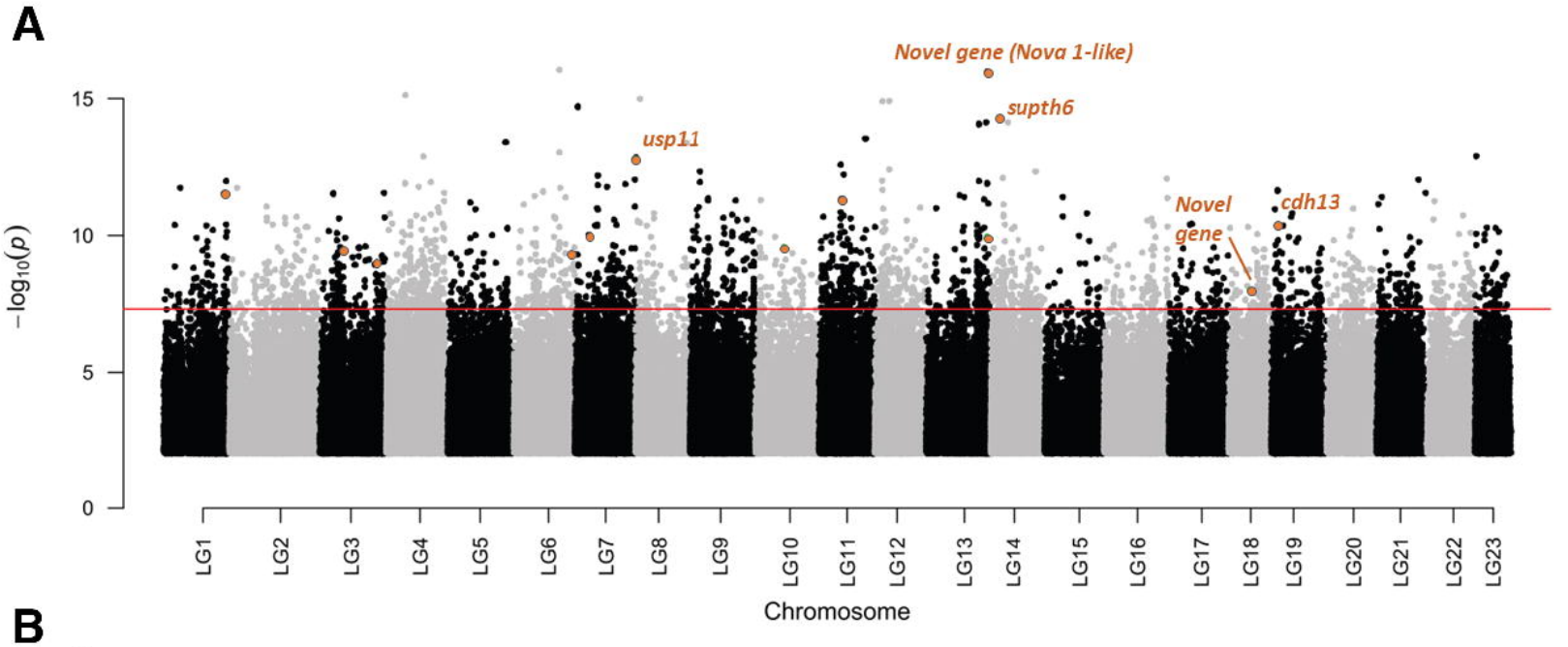
Table S10: Gene Ontology (GO) term of interest associated BFDCA DC gene pairs

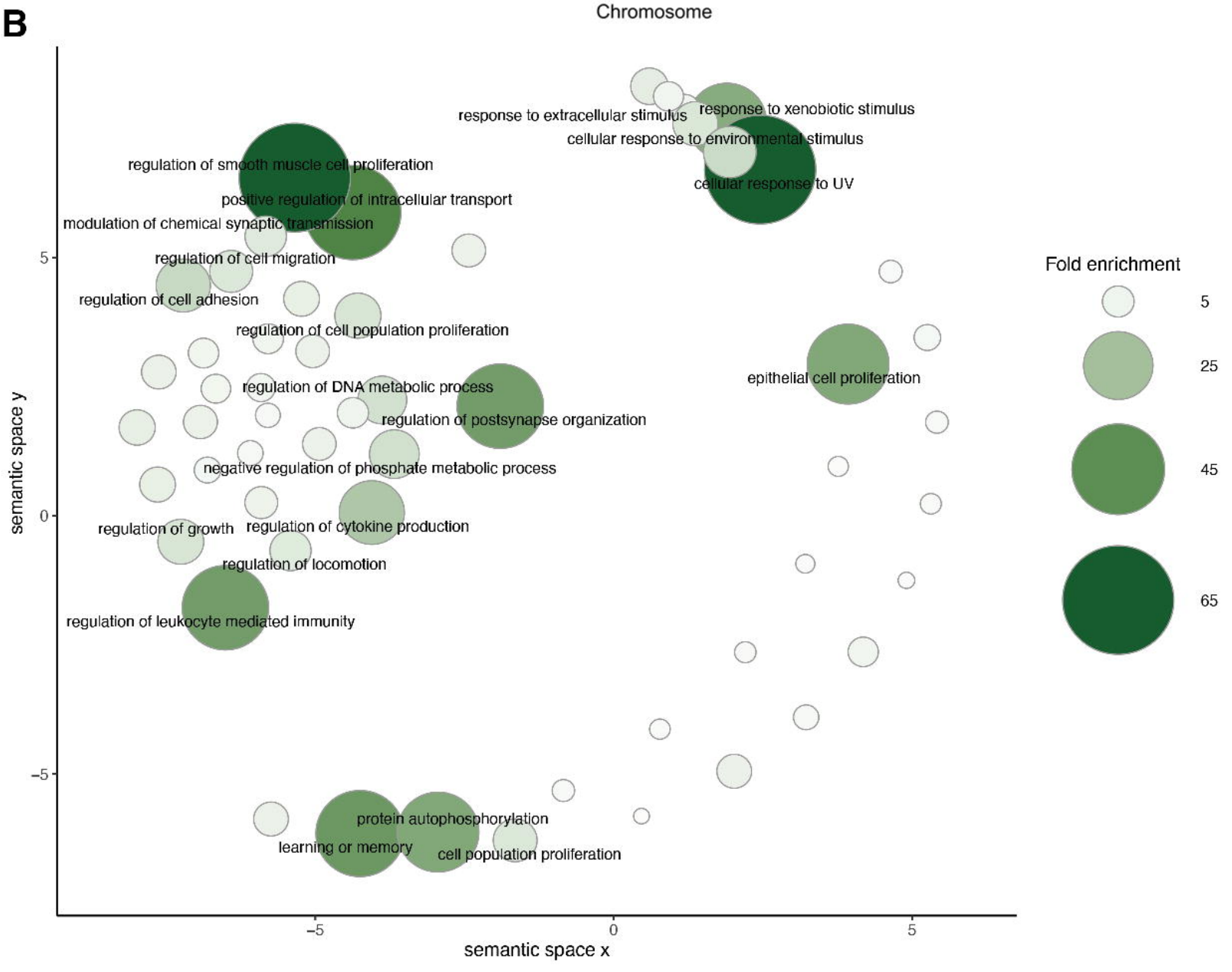
OPTIC TECTUM	
Polarization-selected lines	
GO terms of interest (enriched)	Genes of interest
response to stimulus	TLK2,RTN4,PARVA,PFKFB2,SHC3,TOM1L2,ASIC1,CDH6,GRID2,TSPAN8,CDK14,FOXP2,HDGFL3,NPAS2,NSG2,PHF1,PTPRD,ST3GAL4,UBE2V2,CHD8,CPEB4,GLUL,GNG7,ITPR1,NT5E,PTPRM,SRSF5,TSC22D2,VPS45,XPC
signaling	TLK2,RTN4,PFKFB2,SHC3,TOM1L2,ASIC1,CDH6,GRID2,CDK14,HDGFL3,NSG2,PTPRD,APBA1,CHD8,CPEB4,GLUL,GNG7,ITPR1,NT5E,PTPRM,XPC
response to stress	TLK2,RTN4,ASIC1,TSPAN8,NPAS2,PHF1,ST3GAL4,UBE2V2,CPEB4,GLUL,ITPR1,NT5E,TSC22D2,VPS45,XPC
signal transduction	TLK2,RTN4,SHC3,TOM1L2,ASIC1,CDH6,GRID2,CDK14,HDGFL3,NSG2,PTPRD,CHD8,CPEB4,GNG7,ITPR1,NT5E,PTPRM,XPC
cell-cell signaling	PFKFB2,ASIC1,GRID2,CDK14,PTPRD,APBA1,CHD8,GLUL,ITPR1
response to external stimulus	PARVA,SHC3,GRID2,TSPAN8,ST3GAL4,CHD8,CPEB4,GLUL,NT5E,PTPRM,XPC
synapse	RTN4,ASIC1,GRID2,KLHL17,HAPLN1,PTPRD,APBA1,CPEB4,GLUL,ITPR1,RPL8,VPS45
postsynaptic density	RTN4,GRID2,KLHL17,CPEB4,ITPR1,RPL8
asymmetric synapse	RTN4,GRID2,KLHL17,CPEB4,ITPR1,RPL8
postsynaptic specialization	RTN4,GRID2,KLHL17,CPEB4,ITPR1,RPL8
postsynapse	RTN4,GRID2,KLHL17,APBA1,CPEB4,ITPR1,RPL8
Control lines	
response to stimulus	CGAS,ABCA1,CENPS,STIMATE,BCL11B,DLGAP2,ENAH,MAPK6,TMBIM6,ADH5
positive regulation of cGMP-mediated signaling	CGAS
neurogenesis	BCL11B,ENAH,GBA2,MAPK6
response to L-glutamate	TMBIM6
response to external stimulus	CGAS,ABCA1,BCL11B,ENAH,ADH5
regulation of signal transduction	CGAS,ABCA1,STIMATE,DLGAP2,TMBIM6
neurotransmitter reuptake	SLC29A2
TELENCEPHALON	
Polarization-selected lines	

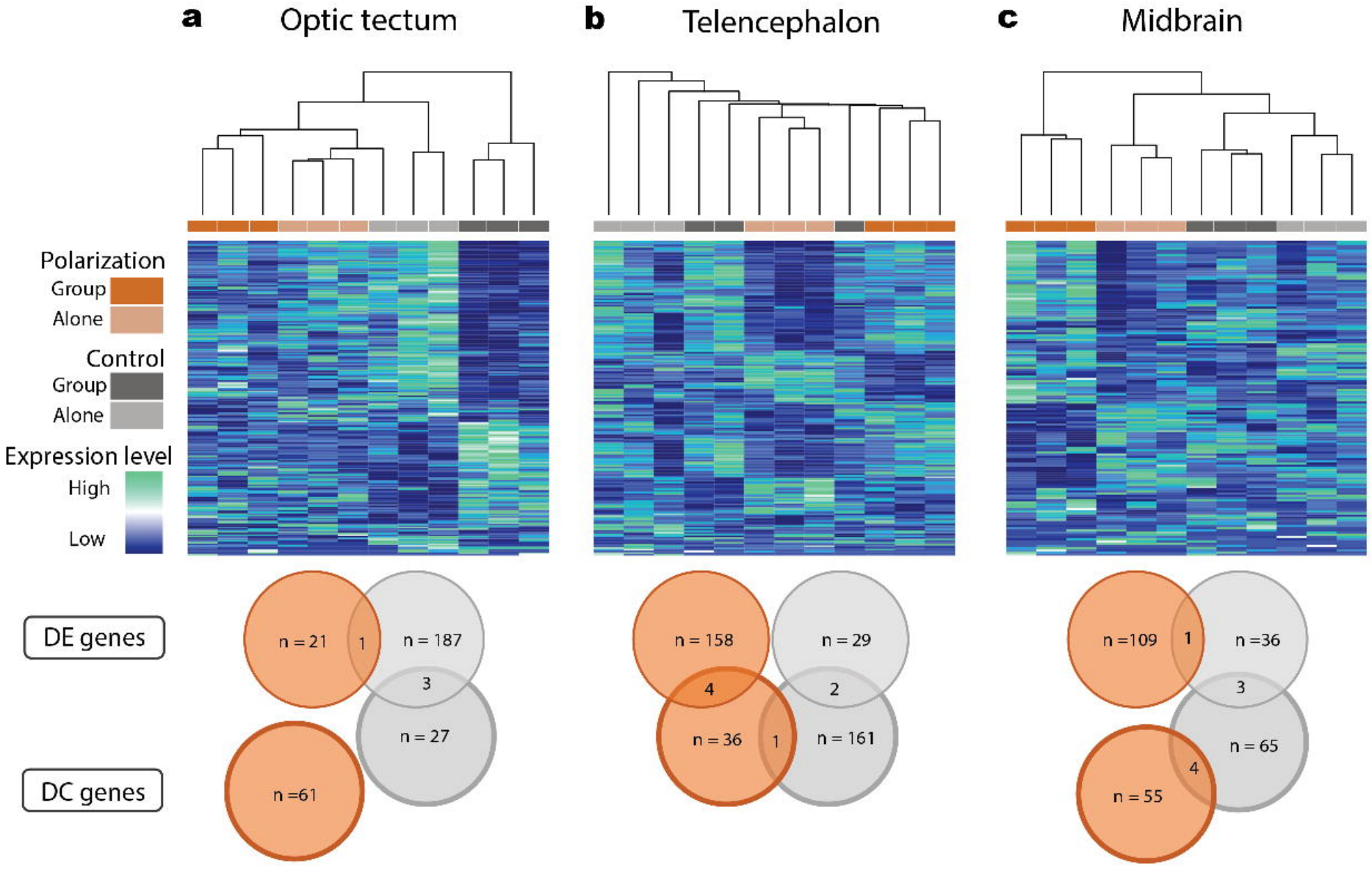
L-glutamate transmembrane transporter activity	SLC17A6, GRM1
neurotransmitter transmembrane transporter activity	SLC17A6 GRM1
glutamate receptor activity	GRM1
postsynaptic neurotransmitter receptor activity	SLC17A6, MACF1, RAB3A,
L-glutamate transmembrane transport	RAB3A, MACF1
evoked neurotransmitter secretion	MAPK9
protein localization to tricellular tight junction	MAPK9
JUN phosphorylation	PTPRS
regulation of cell junction assembly	SLC17A6
neurotransmitter loading into synaptic vesicle	GRM1
adenylate cyclase-inhibiting G protein-coupled glutamate receptor signaling pathway	SLC17A6
neurotransmitter transport	GRM1, MACF1
L-glutamate import across plasma membrane	GRM1
G protein-coupled glutamate receptor signaling pathway	MAPK9
protein localization to cell-cell junction	PTPRS, TENM2, PTPRS, GRM1, MACF1, PNN, RAB3A, RHEB, RPL7, PPP1R13L
cell junction assembly	MTHFS
glutamate metabolic process	GRM1
glutamate receptor signaling pathway	
clathrin-sculpted glutamate transport vesicle membrane	RAB3A RAB3A
clathrin-sculpted glutamate transport vesicle	PNN
exon-exon junction complex	PTPRS
glutamatergic synapse	
Control lines	
high-affinity glutamate transmembrane transporter activity	SLC1A1,SLC1A2 SLC1A1,SLC1A2
glutamate:sodium symporter activity	KAT7,PKN1,SMARCA4,DVL1,GATAD2B,IER2,JARID2,SCAI,TBL1XR1,TNRC6B,AGRN,HDAC4,PMM2,PSMC3,RPL10,TACC1,UBA3,USP9X,CHD5,KDM4A,PPARGC1B,SPEN,UBA52,XPO1,ZNF521,ELAVL2,GPR37,HSPD1,MYT1
aromatic compound biosynthetic process	
synapse organization	DVL1,ACHE,AGRN,BDNF,NTRK2,SLC1A1,GRM5,NLGN1,UNC13A,BSN
glutamate receptor signaling pathway	GRIA2,SLC1A1,GNAQ,GRM5
neurotransmitter transport	DVL1,SLC1A1,NLGN1,SLC1A2,SLC6A11,UNC13A
MIDBRAIN	
Polarization-selected lines	

protein tyrosine kinase activity	PDGFRB,IGF1R
cellular aromatic compound metabolic process	CPSF3,POU4F1,HES6,NUP205,OTP,PDGFRB,CTPS1,
	GAPDHS,MRTFB,POLR1B,RPL19,RPL37A,SFMBT1,
aromatic compound biosynthetic process	SIN3A,TLE4,IL1A,NCOA4,ORC4,PFKFB2
	CPSF3,POU4F1,HES6,OTP,PDGFRB,CTPS1,MRTFB,
inactivation of MAPKK activity	POLR1B,SFMBT1,SIN3A,TLE4,IL1A,NCOA4
signal transduction	IGF1R
	POU4F1,PDGFRB,ARL6,CORO2A,IGF1R,NRG2,TLE
equilibrioception	4,TRIM39,CDC42BPB,IL1A,NCOA4
regulation of signal transduction	POU4F1
suckling behavior	POU4F1,PDGFRB,IGF1R,NRG2,TLE4,TRIM39,IL1A
	POU4F1
Control lines	
signaling receptor activity	CSK,IGF1R,PPARA,GRIA2,NLGN2
signal transduction	MAGI1,CSK,IGF1R,PPARA,HMGA2,RASSF5,DCTN1
behavior	,DOCK9,GRIA2,GRIP1,ITPKB,LBH,NLGN2,NPRL3,P
glutamatergic synapse	RICKLE2,TUFT1,CNGB1,GRK4,MYO1E,RUNDC3A,
AMPA glutamate receptor complex	SCHIP1
	PPARA,DCTN1,NLGN2,ELAVL4
	GRIP1,ATP2B2,ELAVL4
	GRIA2


A
Treatment
Control
Polarization−selected
Females
Males
p = 0.047
p = 0.20
p = 0.043
p = 0.40
Alignment to group direction
Nearest neighbor distance (mm)
Parental
F1
Generation
B
Alignment
Nearest neighbor distance
Same-sex pedigree
Full pedigree
h2[♀]
h2[♂]
r[♀♂]
Estimate ± CI


A
Novel gene (Nova 1-like)
supth6
usp11
Novel gene
cdh13
15
10
5
0
−log10(p)
LG1
LG2
LG3
LG4
LG5
LG6
LG7
LG8
LG9
LG10
LG11
LG12
LG13
LG14
LG15
LG16
LG17
LG18
LG19
LG20
LG21
LG22
LG23
Chromosome


B
response to extracellular stimulus
response to xenobiotic stimulus
cellular response to environmental stimulus
cellular response to UV
regulation of smooth muscle cell proliferation
positive regulation of intracellular transport
modulation of chemical synaptic transmission
regulation of cell migration
regulation of cell adhesion
regulation of cell population proliferation
regulation of DNA metabolic process
regulation of postsynapse organization
negative regulation of phosphate metabolic process
regulation of growth
regulation of cytokine production
regulation of locomotion
regulation of leukocyte mediated immunity
epithelial cell proliferation
protein autophosphorylation
learning or memory
cell population proliferation
Fold enrichment
5
25
45
65
semantic space y
semantic space x
5
0
−5
−5
0
5


a
Optic tectum
b
Telencephalon
c
Midbrain
Polarization
Group
Alone
Control
Group
Alone
Expression level
High
Low
DE genes
DC genes
n = 21
1
n = 187
3
n = 27
n =61
n = 158
n = 29
4
2
n = 36
1
n = 161
n =109
1
n =36
3
n = 65
4
n = 55

DNAseq
RNAseq TEL
RNAseq MBR
RNAseq OT
learning and memory (behavior)
learning and memory (cognition)
protein autophosphorylation
regulation of postsynapse organization
regulation of trans−synaptic signaling
modulation of chemical synaptic transmission
positive regulation of neuron projection development
regulation of DNA metabolic process
negative regulation of phosphate metabolic process
regulation of locomotion
cellular response to environmental stimulus
cellular response to UV
cell population proliferation
regulation of cell population proliferation
positive regulation of intracellular transport
regulation of smooth muscle cell proliferation
regulation of cell adhesion
regulation of cytokine production
regulation of growth
positive regulation of growth
BIOLOGICAL PROCESSES
glutamatergic synapse
neuron to neuron synapse
postsynaptic density
neuronal cell body
apical part of cell
plasma membrane raft
membrane microdomain
cell body
CELLULAR COMPONENTS
identical protein binding
cadherin binding
calcium−dependent protein binding
mRNA 3'−UTR AU−rich region binding
MOLECULAR FUNCTIONS
Fold enrichment
0 25 50 75 100