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], Mirjam Amcoff[2], Severine D. Buechel[8], Maksym Romenskyy[9], Niclas Kolm[2], Judith E. Mank[1]

Affiliations

1 - Department of Zoology, University of British Columbia, Vancouver, Canada

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

3 - Division of Biosciences, University College London, London, United Kingdom

4 – Department of Biomedical Engineering, Los Andes University, 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 Life Sciences, Imperial College London, London, UK.

*corresponding author: corral@zoology.ubc.ca

Abstract

The organization and coordination of fish schools provide a valuable model to investigate the genetic architecture of affiliative behaviors, providing a valuable model to determine the molecular mechanisms underlying social behaviours and personalities. We used quantitative genetic methods to phenotype the sociability of guppy selection lines that vary in group alignment and schooling propensity. We found that the two major traits forming the sociability personality axis in this species showed heritability estimates in the high-range of previous estimates across social behaviors, with important cross sex genetic differences. Additionally, we found that evolution of higher coordinated motion 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. We used PoolSeq and RNASeq to assess the molecular mechanisms underlying sociability in this species and find that regulation of genes involved in neuron migration and synaptic function within the social decision network are key in the evolution of schooling, highlighting a crucial role of glutamatergic synaptic function and calcium-dependent signaling processes.

Introduction

Group-living is a widespread phenomenon across the animal kingdom that 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, is a fundamental aspect of personality, influencing social interactions and often subject to strong selective processes[2,3]. Studies in a range of animals have consistently revealed substantial intra-specific variation in sociability[4–6] , suggesting that individual genetic differences underlie variability in personality and social behavior phenotypes. However, heritability estimates suggest a complex, polygenic architecture [7]. Human twin and family studies suggest that personality traits generally exhibit heritabilites of ≈ 0.4 (reviewed in [8]). In non-human animals, meta-analysis estimated heritability is 0.235 across social behaviors, including personality traits[9,], with heritability of affiliative associations ranging from 0.11 to 0.51 across a range of animals[10–13]. Moreover, behaviour, and in particular sociability, entails indirect genetic effects[14,15], complicating studies of genetic architecture.

Despite this complexity, many neural and molecular mechanisms underlying social behavior have been identified. Many of the neuromodulators (serotonin, dopamine, vasopressin and oxytocin) and neural structures are highly conserved within the social-decision making network across vertebrates[16] . Many human personality traits associated with social-decision making have been linked to dopaminergic and serotonergic genes (reviewed in [17]), and the regulation of these neuromodulators is often connected with neurodevelopmental disorders with strong consequences for affiliative behaviors, such as autism spectrum disorder[6,18]. Studies in non-human organisms likewise point towards a major role of genes involved in the regulation of these neurochemical systems. For instance, sociability is affected in mouse knock-out mutants for genes involved in dopaminergic signalling in the mouse nervous system[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]. 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 social systems.

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[21]. We previously showed that schooling behavior can increase substantially over just three generations of artificial selection in female guppies, *Poecilia reticulata*[22]. and we observed a 15% increase in intrinsic schooling propensity compared to controls. 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 variation to genetic differences, and in this study, we phenotyped the sociability of a large number of guppies (n = 1486) from our experimental selection lines to estimate the heritability of key traits for the sociability axis. Interestingly, because many guppy social interaction patterns have sex differences[23–25], and because our selection was performed solely in females, we are able to examine cross-sex genetic correlations in this ecologically relevant behavioural trait. Genomic and transcriptomic data from these lines reveal the genetic architecture of sociability, highlighting a series of genes with well-defined roles in neurodevelopmental processes and providing a robust agreement across experiments in that genetic regulation of neural processes in decision-making and motor control regions of the brain is central for the 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 how guppies socially interact with unfamiliar conspecifics. For this, we assessed sociability in 740 females and 746 males from multiple families of three polarization-selected and three 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 that capture the most biologically relevant aspects of the sociability axis in this species [21].

Female guppies from polarization-selected lines presented higher alignment and higher attraction to an unfamiliar group compared to control lines ($LMM_{alignment}$: 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 ($LMM_{attraction}$: sex: t = 0.51; df = 447.05; p = 0.61; Fig. 1A; Supplementary Table 1). There were strong 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 Tables 1). Differences in body size between age classes in guppies might have driven these results[26], as the time restrictions involved in testing large numbers of fish required that we assessed parents and offspring 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 at a similar age within parental and offspring generations.

To assess the heritability of sociability in this species, we fitted animal models with alignment and attraction phenotypes quantified from these 1486 individuals. In total, we phenotyped parents, three male and three female offspring for 195 families (99 polarization-selected and 96 control-families). Given known differences between guppy sexes in social interaction patterns[23–25], we estimated heritability with animal models that only included relationships with same sex individuals (same-sex pedigree) or that include relationship with individuals from both sexes (full pedigree).

Using same-sex pedigree animal models, attraction heritability was similar in females ($h^2_{attraction}$ = 0.18; L95% = 0.05; U95% = 0.34; Fig. 1B) and males ($h^2_{alignment}$ = 0.19; L95% = 0.06; U95% = 0.34; Fig. 1B), however alignment heritability was far higher in females ($h^2_{alignment}$ = 0.34; L95% = 0.18; U95% = 0.49; Fig 1B) than males ($h^2_{alignment}$ = 0.06; L95% ≈ 0; U95% = 0.18; Fig. 1B).

Full-pedigree models indicated lower heritability estimates than those obtained from same-sex pedigree models (Supplementary Table 4; Fig 1B), with the exception of the heritability estimate of attraction in males ($h^2_{attraction}$ = 0.26; L95% = 0.16; U95% = 0.37; Fig 1B). In line with this result, we found a positive female-male correlation in attraction (r_{fm} attraction: 0.68; L95% = 0.23; U95% = 0.98; Fig. 1B). Yet, the magnitude of this correlation was difficult to infer from our data given large credible intervals. For alignment, credible intervals for r_{fm} span zero, suggesting no correlation between sexes in this trait (Supplementary Table 4; Fig 1B).

Genetic basis of sociability in guppies

We used sequenced DNA pools (Pool-seq) to identify genome wide differences between polarization-selected female guppies that presented high sociability and those that presented low sociability. For this, we focused on measurements obtained from females in analyses of alignment to an unfamiliar group and pooled the DNA from mothers whose family (normalized mother and daughters alignment score; see methods) were in the top 25% and the lower 25% quartiles from each of the three replicated polarization-selected lines (six total pooled samples with 7 individuals each; Supplementary Figure 1).

DNA reads were aligned to the guppy reference genome (Guppy_female_1.0 + MT, RefSeq accession: GCA_000633615.2) to compare genome wide allele frequency differences between high and low sociability guppies. We ran two independent analyses with these aligned sequences. For our first analysis, we merged sequences from the three replicates with high sociability pooled samples and sequences from three replicates with low sociability pooled samples. We filtered merged sequences to 3,004,974 SNPs (see methods) and performed a Fisher's exact test in Popoolation2[27] to identify SNPs that significantly differed in their allele frequencies between guppies with high and low sociability. Using this methodology, we identify 819 SNPs associated to our sociability phenotype (p-value of the Fisher's exact test -log10(p) > 8; Fig 2A). Most significant differences were found in single physically unlinked SNPs across the genome, consistent with a polygenetic architecture of the trait.

Out of these 819 significantly different SNPs, 421 were located within genes or gene promoter regions of the guppy genome and were used for further functional characterization in association

with Gene Ontology (GO) annotations (273 unique genes). We clustered GO terms based on semantic similarities and found significant overrepresentation of biological process terms related to learning and memory, synaptic functioning, response to stimulus, locomotion and growth (Fig. 2B). We likewise found significant overrepresentation of cadherin and calcium-dependent protein binding annotations (molecular components terms; Supp Figure 2) and glutamatergic synapse annotations (cellular components terms, Supp Figure 3).

Second, we look for consistent differences in allele frequencies between high and low sociability pooled samples in our three replicates by performing the Cochran-Mantel-Haenszel test (cmh-test) in Popoolation2. This stringent analysis identified 13 SNPs from ten different chromosomes with consistent significant differences in allele frequencies across the three replicates (cmh-test p-value <0.01 with false discovery rate correction). Five of these SNPs are located within known coding sequence of the guppy genome, of which three are within well-characterized genes in zebra fish and human homologs with important roles for cognitive function – ubiquitin specific peptidase 11 (*usp11*), *supt6* histone chaperone and transcription elongation factor homolog (*supt6h*), and cadherin 13 (*cdh13*); Fig 2A, Table 1). The other two are classified as novel genes, one of them is matched to an RNA-binding protein Nova-1-like gene, similarly associated to motor function and changes in synaptic function (Fig 2A, Table 1).

Neurogenomic response of schooling in guppies

We used transcriptome sequencing to determine differences in gene expression in multiple brain regions of polarization-selected and control females in response to two different social contexts, swimming alone (the "Alone" condition) or schooling in a group (groups of eight unfamiliar females; the "Group" condition). We focused on three separate brain tissues that control distinct functions. The *optic tectum* is involved in sensory processing of visual signals. The *telencephalon* is implicated in decision making. The *midbrain* is associated with motor function in response to auditory and visual stimuli[28,29]. Together, these three brain tissues contain the main components of the social brain network in fish[30,31].

Differential expression analyses. We identified genes differentially expressed between lines under each treatment condition and in each brain region separately, to determine the neurogenomic

response triggered by schooling in both lines. Gene expression analyses indicate very little overlap in differentially expressed (DE) genes between polarization and control lines (Fig. 3, Supplementary Dataset 1), suggesting that females from different selection lines are activating different transcriptional cascades and biological pathways in response to social context. In polarization lines we found an order of magnitude fewer DE genes in the optic tectum than the other brain components (n=21 for optic tectum, n=158 for telencephalon, n=109 for midbrain, each $p_{adj} < 0.05$). Moreover, in the telencephalon and midbrain, DE genes between Alone and Group treatment in the polarization-selection lines were enriched for GO annotations associated with cognition, memory, learning and social behaviour (Supplementary Dataset 1). We found enrichment for these annotation terms for DE genes expressed in the optic tectum, but not in midbrain or telencephalon, of control lines.

Hierarchical clustering analyses of DE genes showed that females from polarization lines in the Group condition cluster uniquely from polarization females in the Alone condition (Fig. 3, Supplementary Table 5). Similarly, females from polarization-selected lines in the Group condition cluster uniquely from females from control lines under the Group and Alone conditions in both the telencephalon and the midbrain, suggesting a unique response in the regions of the brain associated with behaviour to social exposure. This was not observed in samples from the optic tectum, suggesting that visual processing of social treatments did not differ between polarization-selected and control females. Hierarchical clustering analyses using all expressed genes clustered samples by selection line rather than by social context condition (Supplementary Figure 4), suggesting that social context affects only a targeted subset of the overall transcriptome rather than the majority of genes.

Functional characterization of genes of interest across experiments

We combined the information from our genomic and transcriptomic analyses on polarization-selected and control lines to obtain an intersected delimitation of the gene functions that our analyses highlight as important in the development and expression of social interactions with conspecifics. Specifically, we compared functional analyses in the set of genes differentiated between merged sequences of the three replicates with high and low sociability (273 unique genes) to functional analyses of genes differentially expressed in brain tissue of females following

exposure to multiple social conditions. We find extensive concordance between these functional analyses in independent experiments assessing social interaction of polarization-selected female guppies (Fig. 4). For biological process GO terms, we find an overlap between analyses in genes associated to learning and memory, synaptic processes, neuron projection and cell growth. The overlap between differential SNPs and differential expression is mostly constrained to the telencephalon and midbrain regions (Fig. 4). We found similar patterns in relation to cellular component GO terms, with strong overlap in neuronal components, in particular with high enrichment of terms associated to glutamatergic synapse. Visualization of GO terms associated with molecular functions suggests a major role of genes with protein binding function across experiments, including a role of cadherin-binding related genes in the optic tectum of control lines and the midbrain of polarization-selected lines (Fig. 4).

Discussion

We used behavioural phenotyping across guppy families, in conjunction with PoolSeq and RNASeq to identify the genetic architecture of coordinated motion. Our broad range of analyses, spanning genomes, transcriptomes, and phenotypes, provides an exceptional evaluation of the molecular mechanisms underlying sociability in this fish. Our work suggests that genes and gene networks involved in social decision making through neuron migration and synaptic function are key in the evolution of schooling, highlighting a crucial role of glutamatergic synaptic function and calcium-dependent signaling processes.

Our pedigree-based phenotyping analyses of 195 guppy families from polarization-selected and control lines indicate moderate levels of heritability (Alignment: 0.06 - 0.34; Attraction: 0.09 - 0.26), with pronounced sex differences in full pedigree models ($\text{Alignment}_{\text{female-male}}$: 0.10 ± 0.05; $\text{Attraction}_{\text{female-male}}$: -0.17 ± 0.05), in estimates for key behavioral traits forming the sociability axis in this species. Our heritability estimates are similar to estimates for affiliative social behavior traits in primates, ungulates and rodents[10–13,32], and to overall estimations of heritability in personality traits across human and non-human animals[8,33]. Overall, our heritability measures indicate a complex genetic architecture, as expected for a behavioural trait with indirect genetic effects[14]. Given the importance of social behaviour in a range of survival and fitness components

in natural systems[1,58,59], our results suggest that complex genetic architectures can respond to strong evolutionary pressures, even when only one sex is subject to selection. [57].

We found a positive and significant cross sex genetic difference for attraction (0.68, CI: 0.25 - 0.98), suggesting genetic differentiation in the genes acting on this trait between males and females. In contrast, cross sex genetic difference estimates were not significant (0.26, CI: -0.58 - 0.92). Yet, the observed sex differences in the heritability estimates of alignment and cross sex genetic differences in attraction should be expected to facilitate the evolution of sexually dimorphic for sociability in this species. To our knowledge, only one previous study has assessed cross sex genetic differences in a sociability proxy in non-human animals, showing that genetic architecture of group size choice is decoupled in fruit flies[34]. Overall, there are few studies evaluating the interaction between genotype and sex for other personality traits in non-human animals, and most are focused on the bold-shy continuum aspect of personality where evidence for the presence and the direction of sex-specific genetic variance is mixed[35–37]. In guppies, sex differences in risk-taking behaviors are weak and likely lack sex-specific genetic architecture[37]. Our results here add to a growing body of work of sex-differences in the genetic architecture of behavioural traits, and suggest that future studies should account for sex-specific additive genetic variance when inferring the evolvability of personality traits.

We next mapped the genomic and transcriptomic basis of phenotypic differences in polarization in female guppies. Our most stringent PoolSeq analysis identified SNPs in four genes that consistently differed across all three replicate lines. s*upt6* histone chaperone and transcription elongation factor homolog (*supt6h*) is important in the positive regulation of transcriptional elongation and a substrate of mTOR, a signalling pathway with a role in cognitive function[38,39]. Ubiquitin specific peptidase 11 (*usp11*) homolog in humans has a critical function in the development of neural cortex, and knock-out studies in mice show that the locus protects females from cognitive impairment[40,41]. Similarly, the novel RNA-binding protein Nova-1-like gene is associated with a neuron-specific nuclear RNA binding protein in humans and regulates brain-specific splicing related to synaptic function (Buckanovich, Ule et al. 2005).

Finally, our PoolSeq analysis identified Cadherin 13 (*cdh13*), the human homolog of which has a crucial role in GABAergic function[42], with involvement in neural growth and axonal guidance during early development[43,44]. Moreover, deficit of this gene has a major impact in neurodevelopmental disorders including attention-deficit/hyperactivity disorder and autism spectrum disorder[45]. Indeed, *cdh13* knock-our mice display delayed acquisition of learning tasks and a decreased latency in sociability[46] . Interestingly, repeated selection of genes involved in cadherin-signalling pathways[47] has been shown in guppy populations experiencing different predation pressures. Together, natural selection imposed by differences in predation across these populations[48–50] , and our combined findings in the genomic background of guppies suggest strong selective pressures for cadherin-signalling genes due to their modulation of affiliative behaviors.

Our GWAS experimental design was built around comparing individuals with high- and low-sociability phenotypes from within polarization-selected lines, rather than between polarization-selected and control lines. This may have resulted in compressed phenotypic spread, but carries the important advantage of reducing the incidence of SNP frequencies that vary across alternative selection lines due to drift. As such, our design is conservative.

Our expression results revealed differences in regulation in genes associated with learning, behavior and neural function, mainly in the telencephalon and midbrain, in comparisons of polarization-selected and control lines in different social contexts. Overall, this suggests that the regulation of highly demanding cognitive processes via synaptic function underlies variation in sociability. While the integration of visual signals is central in fish schools[51], our results suggest that higher order cognitive process are the basis of variation in social affinity. Indeed, the differences in alignment and attraction observed when swimming with unfamiliar conspecifics are arguably highly cognitively demanding, as within a collective motion context, the tendency to copy others' directional movements implies a direct trade-off between personal goal-oriented behaviors and the benefits of social conformity[52,53]. Moreover, variation in polarization is associated with other cognitively demanding behaviors, such as social information use for anti-predator behavior[54]. Together, our study of transcriptomic profiles of schooling fish suggests that the regulation of affiliative behaviors of this species is driven by an intricately linked social

decision-making network in the brain[30], with strong links to functional groups governing social behaviours and personality across species.

We find a striking functional concordance in the functionality of genes identified in genomic and transcriptomic profiles from independent experiments assessing social interactions of polarization-selected female guppies (Fig. 4). The concordance in learning, synaptic processes and neuron projection Gene Ontology terms restricted to brain regions associated with decision-making and motor control, strongly reinforces the notion that genetic regulation of these cognitive processes is fundamental for sociability in the species. Our results indicate that the regulation of glutamatergic synaptic processes is a particularly promising network for future studies of affiliative behaviors. Interestingly, differential gene expression of glutamate receptor genes has been identified to independently evolved across species of vertebrates for the regulation of long-term affiliative behaviors in a mating context[55]. Our results are also concordant with other comparative transcriptomic studies of behavioral responses towards conspecific territorial intrusions, which identified calcium ion-binding regulation across phylogenetically distant species[56]. Together, the consistency in our findings of specific genes and functional terms associated with calcium-dependent and cadherin binding molecular functions across our experiments suggest these are promising molecular targets for future research exploring the evolution and regulation of sociability and affiliative behaviors.

Methods

Study system

To evaluate the genetic architecture of sociability, we performed a series of experiments in guppies following experimental evolution of coordinated motion in this species. The artificial selection procedure is outlined in detail in [22,60]. In brief, groups of female guppies were subjected to repeated open field tests, and were subsequently sorted based on their median polarization, measured by the degree of alignment exhibited by the individuals within the group when swimming together[22,57,61]. For three generations, females from groups with higher polarization were mated with males from those cohorts to generate three lines of guppies that had been selected for high polarization. In parallel, random females were exposed to the same experimental conditions and were mated with

unselected males to generate three control lines. Analysis of the third generation of polarization-selection revealed that, on average, females exhibited a 15% higher level of polarization and a 10% higher level of group cohesiveness compared to control females[57].

Throughout the selection experiment and the completion of experiments described below all fish were removed from their parental tanks after birth, separated by sex at the first onset of sexual maturation, and afterwards kept in single-sex groups of eight individuals in 7 L tanks containing 2 cm of gravel with continuously aerated water, a biological filter, and plants for environmental enrichment. We allowed for visual contact between the tanks. The laboratory was maintained at 26°C with a 12-h light:12-h dark schedule. Fish were fed a diet of flake food and freshly hatched brine shrimp daily.

Social interactions with unfamiliar individuals and the heritability of sociability

To investigate the heritability and cross-sex genetic correlations of social behaviors in the guppy, we measured social interactions of parents and offspring from polarization-selected and control lines with unfamiliar groups of conspecifics. Using offspring of the F3 generation of selection we bred 35 families for each of the three polarization-selected and for each of the three control lines. Specifically, from our population of F3 generation offspring (kept in single sex groups prior to the breeding experiments), we used male and female guppies of the same age (approximately 9 months old) and paired them in 3L tanks to generate the parental generation. We collected offspring for the two first clutches of these pairs and we transferred newborn offspring to 3L tanks in groups of three siblings. We separated siblings by sex at the first onset of sexual maturation, and afterwards we kept them in single-sex groups of three individuals until behavioural testing. We phenotyped sociability of a total of 195 guppy families: mother, father and six offspring (three females and three males). Any family for which we did not collect at least three female and three male offspring was disregarded from further behavioral testing. Each of the six selection lines was represented with a minimum of 30 families in our heritability analyses.

Behavioral assays. To phenotype sociability in guppy families, we measured social interactions of 1495 guppies from our breeding experiment. For each fish, we performed an open field assay using white arenas with 55 cm diameter and 3 cm water depth in which our focal fish (guppies

from the breeding experiment) interacted with a group of seven same-sex conspecifics. Non-focal guppies used in these assays were from a lab wild-type stock population and of similar age to our focal fish. Prior to the start of the test, focal fish and the seven-fish group were confined in the centre of the arena for one minute in separate opaque white 15 cm PVC cylinders. After this acclimation period, we lifted the cylinders and filmed the arena for 10 minutes using a Point Grey Grasshopper 3 camera (FLIR Systems; resolution, 2048 pixels by 2048 pixels; frame rate, 25 Hz). Three weeks prior to assays, we tagged non-focal fish with small black elastomere implants (Northwest Marine technology) to allow recognition of focal fish after completion of each assay. After completion, we gently euthanized focal fish from the parental generation with an overdose of benzocaine and kept them in ethanol for future genomic analyses. Focal fish from the offspring generation were transferred to group tanks for future experimental use. Groups of seven-fish wild-type fish were transferred to holding tanks and used in a maximum of 7 assays with focal fish

Data processing. We tracked the movement of fish groups in the collected video recordings using IDTracker[62] and used fine-grained tracking data to calculate the following variables in Matlab 2020 following methods established in [22]: i) alignment, the median alignment of the focal fish to the group average direction across all frames in the assay; ii) attraction, the median nearest neighbor distance across all frames in the assay; and iii) activity, the median speed frames in each assay. We calculated these variables for the focal fish and the average for the seven-fish wild-type group. Calculations for alignment only considered frames in which at least six individuals formed a connected group, with an interindividual distance of less than 10cm counting as a connection. For all variables, we disregarded tracking data that did not present a minimum of 16 consecutive tracked frames.

Statistical analyses. We used Linear Mixed Models (LMMs) with alignment and attraction as dependent variables to test for potential differences between polarization-selected and control lines in social interactions with unfamiliar individuals. Selection regime, sex, the interaction between these two factors and generation were included as fixed effects. The average activity of the wild-type group was coded as a covariate, with a random intercept for each replicated selection line, the breeding family, and the number of tests previously performed with the wild-type group as random factors. All models were run in R (v.4.1.3) using lme4 and lmerTest packages[63,64]. Model

diagnostics showed that residual distributions were roughly normal with no evidence of heteroscedasticity.

To estimate heritability, the degree of phenotypic variation due to genetic inheritance, and cross-sex genetic correlations of alignment and attraction we used Bayesian animal models[65]. Animal models use a matrix of pedigree relationships set as a random effect, to separate phenotypic variance for each response variable into additive genetic variance and the remaining variance (e.g. maternal effects, environmental effects). Given strong sex differences in social interactions in guppies, we performed three animal models for each trait: one including the full pedigree of the 1495 phenotyped individuals, and two including only the pedigree of phenotyped females or males. Parameter values were estimated using the brms interface[66,67] to the probabilistic programming language Stan[68]. We used normal priors with a mean of 0 and standard deviation of 3 for fixed effects, and student-t priors with 5 degrees of freedom, a mean of 0 and sd of 5 for random effects. The full pedigree model estimated cross-sex correlations with a Lewandowski-Kurowicka-Joe (LKJ) prior with $\eta = 1$, which is uniform over the range -1 to 1. Posterior distributions for full / same-sex pedigree models were obtained using Stan's no-U-turn HMC with 24 / 16 independent Markov chains of 2500 / 4000 iterations, discarding the first 1500 / 2000 iterations per chain as warm-up and resulting in 24000 / 32000 posterior samples overall. Convergence of the chains and sufficient sampling of posterior distributions were confirmed by a potential-scale-reduction metric (R) below 1.01, and an effective sample size of at least 1000. For each model, posterior samples were summarized on the basis of the Bayesian point estimate (median) and posterior uncertainty intervals by HDIs. We calculated estimates of heritability by taking the ratio of the additive genetic variance to the total phenotypic variance in each independent model (see Supplementary Tables 3-4).

Genetic basis of sociability in guppies

Pooled DNA sequencing. For each of the three replicates, we pooled tissue from seven individuals representing the top and bottom 20% polarization-selected female guppies from the parental generation whose families presented higher and lower sociability (Supp Figure 1). Each sample pool was homogenized and DNA was extracted using Qiagen's DNAeasy kits following standard manufacturer's protocol with an additional on-column RNase A treatment. DNA with high

molecular weight from each of 50 individuals per population was quantified using fluorometry (Qubit; Thermo Scientific) and pooled at equal concentrations to achieve 3 µg pooled genomic DNA in a volume in the range of 57–136 µl. We used Nextera DNA Flex library preparation kit following the manufacturer protocol and obtained sequences from SciLife Lab, Uppsala (Sweden). All pools were sequenced in one lane on an Illumina NovaSeq 6000. We obtained on average 31.8 million 150bp read pairs per sample (26.9 million read pairs minimum per sample).

Read quality control and trimming. We assessed the quality of reads for each pool using FastQC v.0.11.4. (www.bioinformatics.babraham.ac.uk/projects/fastqc). After verifying initial read quality, reads were trimmed with Trimmomatic v0.35[69]. We filtered adaptor sequences and trimmed reads if the sliding window average Phred score over four bases was <15 or if the leading/trailing bases had a Phred score <4, removing reads post filtering if either read pair was <50 bases in length. Quality was verified after trimming with FastQC.

Genome-wide allele frequency analysis. We mapped DNA Pool-seq reads against the Ensembl female guppy genome assembly (Guppy_female_1.0 + MT, RefSeq accession: GCA_000633615.2)[70], using HiSat 2.0.5[71]. Each pool's reads were individually mapped to the genome and built into transcripts using default parameters. Reads were mapped using default settings to the guppy reference genome assembly with bwa-mem (v0.7.17). We used Samtools (v. 1.6.0)[72] to convert sam to bam files, to sort bam files, to remove duplicates, and to make mpileup files. First, to identify SNPs that significantly differed in their allele frequencies between guppies with high and low sociability, we merged sequences from high sociability and low sociability pools and used Popoolation2[27] to create a synchronized file with allele frequencies for high and low sociability (mpileup2sync.pl –min-qual 20), compute allele frequency differences (mpileup2sync.pl --min-count 6 --min-coverage 25 --max-coverage 200), calculate Fst for every SNP (fst-sliding.pl) and perform a Fisher's exact test (fisher-test.pl). Second, we similarly used Popoolation2 to detect consistent changes in allele frequencies of sociability pooled samples for our three replicated artificial selection lines. For this, we created one sync file per replicate (mpileup2sync.pl –min-qual 20), and perform a Cochran-Mantel-Haenszel (CMH) test (cmh-test.pl --min-count 18 --min-coverage 25 --max-coverage 200). Using package qqman[73] in R (vs.

4.1.3) we made Manhattan plots for each chromosome by plotting the negative log10-transformed p-values of exact fisher and CMH tests as a function of chromosome position.

Significance tests and functional analyses. We determine SNPs that were significantly different between high and low sociability merged pools in Fisher exact tests using the traditional genome-wide significance threshold [$-\log10(p) > 8$]. We next used custom scripts to identify the overlap between position of these SNPs and genes present in the guppy reference annotated genome[70], and to find homologous genes of these set in medaka (*Oryzias latipes*). We further used this set of unique genes (n = 160) to determine associated GO terms between our merged pools. For this, we performed enrichment tests in PANTHER[74], as implemented in the GO Ontology Consortium (http://www.geneontology.org/). To test for enrichments of GO terms we performed one-tail Fisher's exact tests with a Bonferroni corrected p-value threshold of $p < 0.05$ using a full list of medaka genes orthologous to guppy genes as background. We used Revigo (http://revigo.irb.hr)[75] to find and visualize representative subsets of terms based on semantic similarity measurements for our enriched GO terms related to biological processes, cellular components and molecular functions.

For CMH test results, we determined SNPs that were significantly different between high and low sociability pools based on FDR corrected p-value < 0.05 We used a custom script to identify the overlap between position of these SNPs and genes present in the guppy reference annotated genome.

Neurogenomic response of schooling in guppies

Behavioral assays and tissue collection. We placed individual or groups of eight unfamiliar adult control and polarization-selected females in white 55cm arenas. After 30 min females were euthanized by transfer to ice water. After 30 seconds, with aid of a Leica S4E microscope, we removed the top of the skull and after cutting transversally posterior of the optic tectum and anterior of the cerebellum, and horizontally through the optic chiasm, removed the brain from the skull and placed it into ice water. We severed the *telencephalon* from the rest of the brain between the ventral telencephalon and thalamus at the 'commissura anterioris', including both the pallium and subpallium regions. Then we cut the laminated cup-like structures of the *optic tectum*. The

remaining part of the brain was the *midbrain*. Dissections took under 2 minutes and tissue samples were immediately preserved in RNAlater (Ambion) at 4°C for 24 hours and then at -20°C until RNA extraction.

RNA extraction and sequencing. For each treatment, we pooled tissue from ten individuals into two non-overlapping pools of five for each replicate line in order to obtain sufficient RNA for sequencing. This represents a total of 120 individual females, representing six pools per treatment, per selection regime for a total of 24 pools per tissue. Each sample pool was homogenized and RNA was extracted using Qiagen's RNAeasy kits following standard manufacturer's protocol. Libraries for each sample were prepared and sequenced by the Wellcome Trust Center for Human Genetics at the University of Oxford, UK. All samples were sequenced across 9 lanes on an Illumina HiSeq 4000. We obtained on average 33.9 million 75bp read pairs per sample (28.9 million read pairs minimum, 39.8 million maximum).

Read quality control and trimming. We assessed the quality of reads for each sample using FastQC v.0.11.4. (www.bioinformatics.babraham.ac.uk/projects/fastqc). After verifying initial read quality, reads were trimmed with Trimmomatic v0.35[69]. We filtered adaptor sequences and trimmed reads if the sliding window average Phred score over four bases was < 15 or if the leading/trailing bases had a Phred score < 3, removing reads post filtering if either read pair was < 33 bases in length. Quality was verified after trimming with FastQC.

Differential Expression Analysis. We mapped RNAseq reads against the latest release of the published guppy genome assembly (Guppy_female_1.0 + MT, RefSeq accession: GCA_000633615.2)[70], using the HiSat 2.0.5 – Stringtie v1.3.2 suite[71]. Each pool's reads were individually mapped to the genome and built into transcripts using default parameters. The resulting individual assemblies were then merged into a single, non-redundant assembly using the built-in StringTie-merge function. We filtered the resulting assembly for non-coding RNA using medaka and Amazon molly (*Poecilia formosa*) non-coding RNA sequences as reference in a nucleotide BLAST (Blastn). After eliminating all sequences matching non-coding RNAs, we kept only the longest isoform representative for each transcript for further analysis. Finally, we quantified expression by re-mapping reads to this filtered assembly using RSEM (v1.2.20)[76]

Lowly expressed genes were removed by filtering transcripts < 2 RPKM (reads per kilobase per million mapped reads), preserving only those transcripts that have expression above this threshold in a least half of the samples for each treatment within a line. After this final filter, a total of 26,140 optic tectum transcripts, 25,100 telencephalon transcripts and 26,514 midbrain transcripts were retained for further analysis. Using sample correlations in combination with MDS plots based on all expressed transcripts, we determined that none of the 72 pools represented outliers and thus, all samples were included in the analysis.

We used DESeq2[77] to normalize filtered read counts using standard function identify differentially expressed (DE) genes between the alone and the group treatment in Control and Polarization lines separately, and then examined the overlap in differentially expressed genes between them. A transcript was considered differentially expressed if it had an FDR corrected p-value < 0.05. As behaviour could be modulated by small changes in expression, we did not filter differentially expressed genes based on Log Fold-change in expression between the treatments.

Functional Analyses. To investigate the function of differentially expressed genes we performed a Biological Process GO term and pathway enrichment tests. To accomplish this, we initially completed the annotation of the reference genome assembly. The transcripts without clear gene names from the reference genome, and the de novo transcripts identified by HiSat were annotated with blastX against the Swissprot non-redundant database. We then determined which GO terms and known pathways are associated with differentially expressed genes and performed enrichment tests in PANTHER[74]. To tests for enrichments of GO terms and pathways we performed one-tail Fisher's exact tests with a Bonferroni corrected p-value threshold of $p < 0.05$ using medaka genes orthologous to guppy genes as background.

Acknowledgements

We thank David Sumpter, Kristiaan Pelckmans and James Herbert-Read for important contributions to the conceptualization of the artificial selection procedure. We are grateful to Jacelyn Shu for the design of guppy graphics for figures. We thank A. Rennie, E. Trejo, and A.

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 behavioural 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**: All data and code needed to evaluate the conclusions in the paper will be deposited in an online repository. Genomic and transcriptomic data will be deposited at NCBI. 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.* Brain size does not impact shoaling dynamics in unfamiliar groups of guppies (Poecilia reticulata). *Behavioural Processes* **147**, 13–20 (2018).
23. Piyapong, C. *et al.* Sex matters: a social context to boldness in guppies (Poecilia reticulata). *Behavioral Ecology* **21**, 3–8 (2010).
24. 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.

25. Griffiths, S. W. & Magurran, A. E. Sex and schooling behaviour in the Trinidadian guppy. *Anim Behav* **56**, 689–693 (1998).
26. Hoare, D. J., Krause, J., Peuhkuri, N. & Godin, J.-G. J. Body size and shoaling in fish. *J Fish Biol* **57**, 1351–1366 (2000).
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. Northcutt, R. G. Forebrain evolution in bony fishes. *Brain Res Bull* **75**, 191–205 (2008).
29. Salas, C. *et al.* Neuropsychology of Learning and Memory in Teleost Fish. *Zebrafish* **3**, 157–171 (2006).
30. Bshary, R., Gingins, S. & Vail, A. L. Social cognition in fishes. *Trends Cogn Sci* **18**, 465–471 (2014).
31. Tunstrøm, K. *et al.* Collective States, Multistability and Transitional Behavior in Schooling Fish. *PLoS Comput Biol* **9**, e1002915 (2013).
32. Albery, G. F. *et al.* Multiple spatial behaviours govern social network positions in a wild ungulate. *Ecol Lett* **24**, 676–686 (2021).
33. 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).
34. Geiger, A. P. & Saltz, J. B. Strong and weak cross-sex correlations govern the quantitative-genetic architecture of social group choice in Drosophila melanogaster. *Evolution (N Y)* **74**, 145–155 (2020).
35. van Oers, K., Drent, P. J., de Jong, G. & van Noordwijk, A. J. Additive and nonadditive genetic variation in avian personality traits. *Heredity 2004 93:5* **93**, 496–503 (2004).
36. Han, C. S. & Dingemanse, N. J. Sex-dependent expression of behavioural genetic architectures and the evolution of sexual dimorphism. *Proceedings of the Royal Society B: Biological Sciences* **284**, (2017).
37. 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).
38. Lipton, J. O. & Sahin, M. The Neurology of mTOR. *Neuron* **84**, 275–291 (2014).
39. 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).
40. Chiang, S. Y. *et al.* Usp11 controls cortical neurogenesis and neuronal migration through Sox11 stabilization. *Sci Adv* **7**, 6093–6105 (2021).
41. Yan, Y. *et al.* X-linked ubiquitin-specific peptidase 11 increases tauopathy vulnerability in women. *Cell* **185**, 3913-3930.e19 (2022).
42. 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).
43. Ciatto, C. *et al.* T-cadherin structures reveal a novel adhesive binding mechanism. *Nature Structural & Molecular Biology 2010 17:3* **17**, 339–347 (2010).
44. Hayano, Y. *et al.* The role of T-cadherin in axonal pathway formation in neocortical circuits. *Development* **141**, 4784–4793 (2014).

45. 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).
46. 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).
47. 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).
48. Seghers, B. H. Schooling Behavior in the Guppy (Poecilia reticulata): An Evolutionary Response to Predation. *Evolution (N Y)* **28**, 486 (1974).
49. 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).
50. 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**, (2017).
51. Strandburg-Peshkin, A. *et al.* Visual sensory networks and effective information transfer in animal groups. *Current biology* **23**, R711 (2013).
52. 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).
53. 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).
54. Corral-Lopez, A. *et al.* Evolution of schooling propensity in the guppy drives changes in anti-predator behavior that are linked to neuroanatomy. *bioRxiv* 2022.09.07.506928 (2022).
55. Young, R. L. *et al.* Conserved transcriptomic profiles underpin monogamy across vertebrates. *Proc Natl Acad Sci U S A* **116**, 1331–1336 (2019).
56. Rittschof, C. C. *et al.* Neuromolecular responses to social challenge: Common mechanisms across mouse, stickleback fish, and honey bee. *Proc Natl Acad Sci U S A* **111**, 17929–17934 (2014).
57. Kotrschal, A. *et al.* Rapid evolution of coordinated and collective movement in response to artificial selection. *Sci Adv* **6**, eaba3148 (2020).
58. 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).
59. Parrish, J. K. & Edelstein-Keshet, L. Complexity, Pattern, and Evolutionary Trade-Offs in Animal Aggregation. *Science (1979)* **284**, 99–101 (1999).
60. 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).
61. Szorkovszky, A. *et al.* Assortative interactions revealed by sorting of animal groups. *Anim Behav* **142**, 165–179 (2018).
62. Pérez-Escudero, A., Vicente-Page, J., Hinz, R. C., Arganda, S. & de Polavieja, G. G. idTracker: tracking individuals in a group by automatic identification of unmarked animals. *Nature Methods 2014 11:7* **11**, 743–748 (2014).
63. Bates, D., Sarkar, D., Bates, M. D. & Matrix, L. The lme4 Package. *R package version* **2**, 74 (2007).

64. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest Package: Tests in Linear Mixed Effects Models. *J Stat Softw* **82**, 1–26 (2017).
65. de Villemereuil, P. On the relevance of Bayesian statistics and MCMC for animal models. *Animal Breeding Genetics* **136**, 339–340 (2019).
66. Bürkner, P. C. Advanced Bayesian multilevel modeling with the R package brms. *R J* **10**, 395–411 (2018).
67. Bürkner, P. C. brms: An R package for Bayesian multilevel models using Stan. *J Stat Softw* **80**, 1–28 (2017).
68. Carpenter, B. *et al.* Stan: A probabilistic programming language. *J Stat Softw* **76**, (2017).
69. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
70. Künstner, A. *et al.* The Genome of the Trinidadian Guppy, Poecilia reticulata, and Variation in the Guanapo Population. *PLoS One* **11**, e0169087 (2016).
71. Pertea, M., Kim, D., Pertea, G. M., Leek, J. T. & Salzberg, S. L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature Protocols 2016 11:9* **11**, 1650–1667 (2016).
72. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
73. Turner, S. D. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *J Open Source Softw* **3**, 731 (2018).
74. 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).
75. Supek, F., Bošnjak, M., Škunca, N. & Šmuc, T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* **6**, (2011).
76. 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).
77. 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).
78. 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).
79. 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).
80. Takeichi, M. The cadherin superfamily in neuronal connections and interactions. *Nature Reviews Neuroscience 2006 8:1* **8**, 11–20 (2006).

Figures and Tables

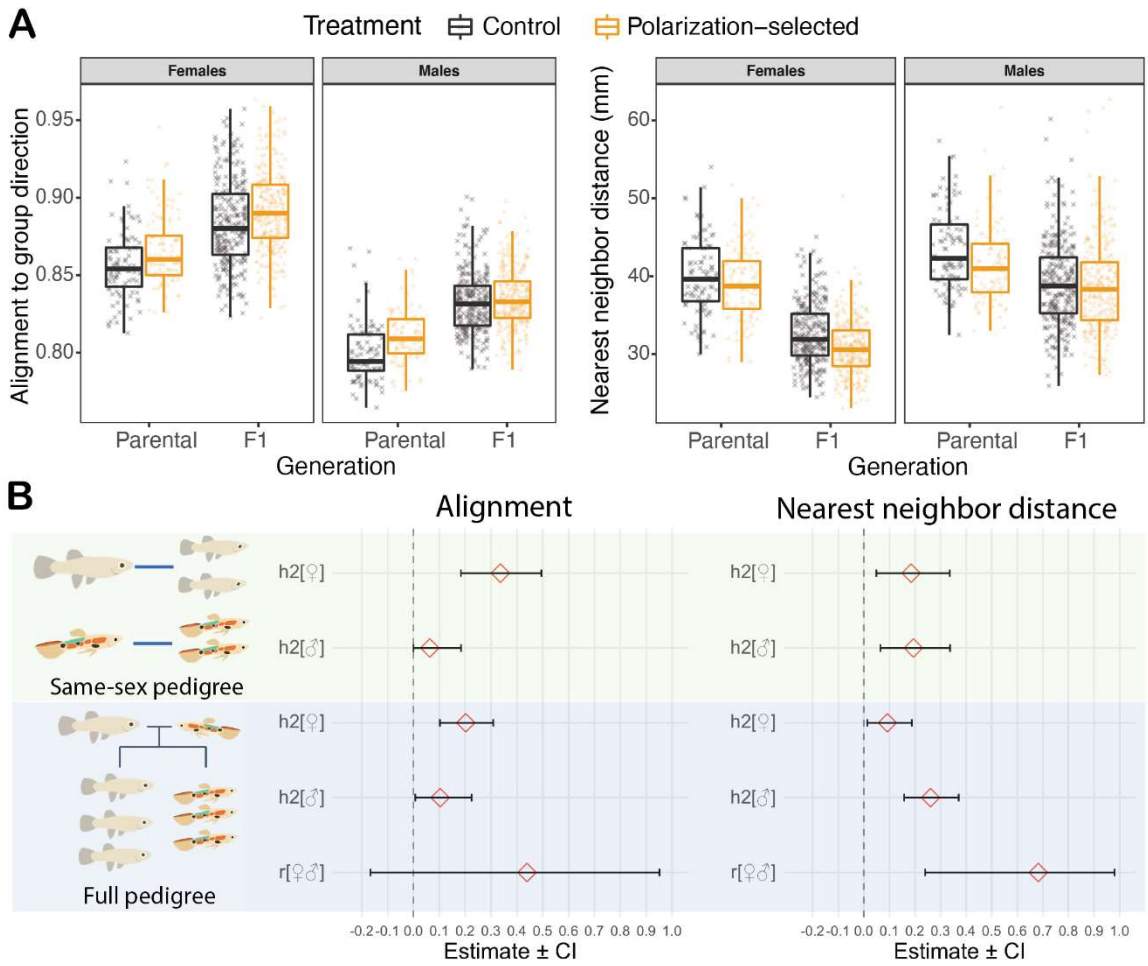

Figure 1. Heritability of sociability in guppies. (A) Guppies from polarization-selected lines (orange) presented higher alignment to the group direction (left panel) and shorter distances to their nearest neighbor (higher alignment; right panel) than guppies from control lines (gray) when swimming with unfamiliar same-sex conspecifics (see Supplementary Table 1). For all boxplots, horizontal lines indicate medians, boxes indicate the interquartile range, and whiskers indicate all points within 1.5 times the interquartile range. **(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 3-4).

Table 1. Characterization of genes associated to sociability in guppies

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 to neurological disorders.	Chiang et al, 2021[40] Karaca et al. 2015[78]
Chr 13: 31383940	00000018946	*Novel gene (RNA-binding protein Nova-1-like)*	Neuronal RNA binding protein associated to motor function.	Buckanovich and Darnell 1997[79]
Chr 14: 4286109	00000009725	*supt6h*	Substrate of mTOR, a signaling pathway associated to brain function and neurodegenerative disorders.	Lipton and Sahin 2015[38] Zhao et al. 2015[39]
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[42] Takaechi 2006[80]

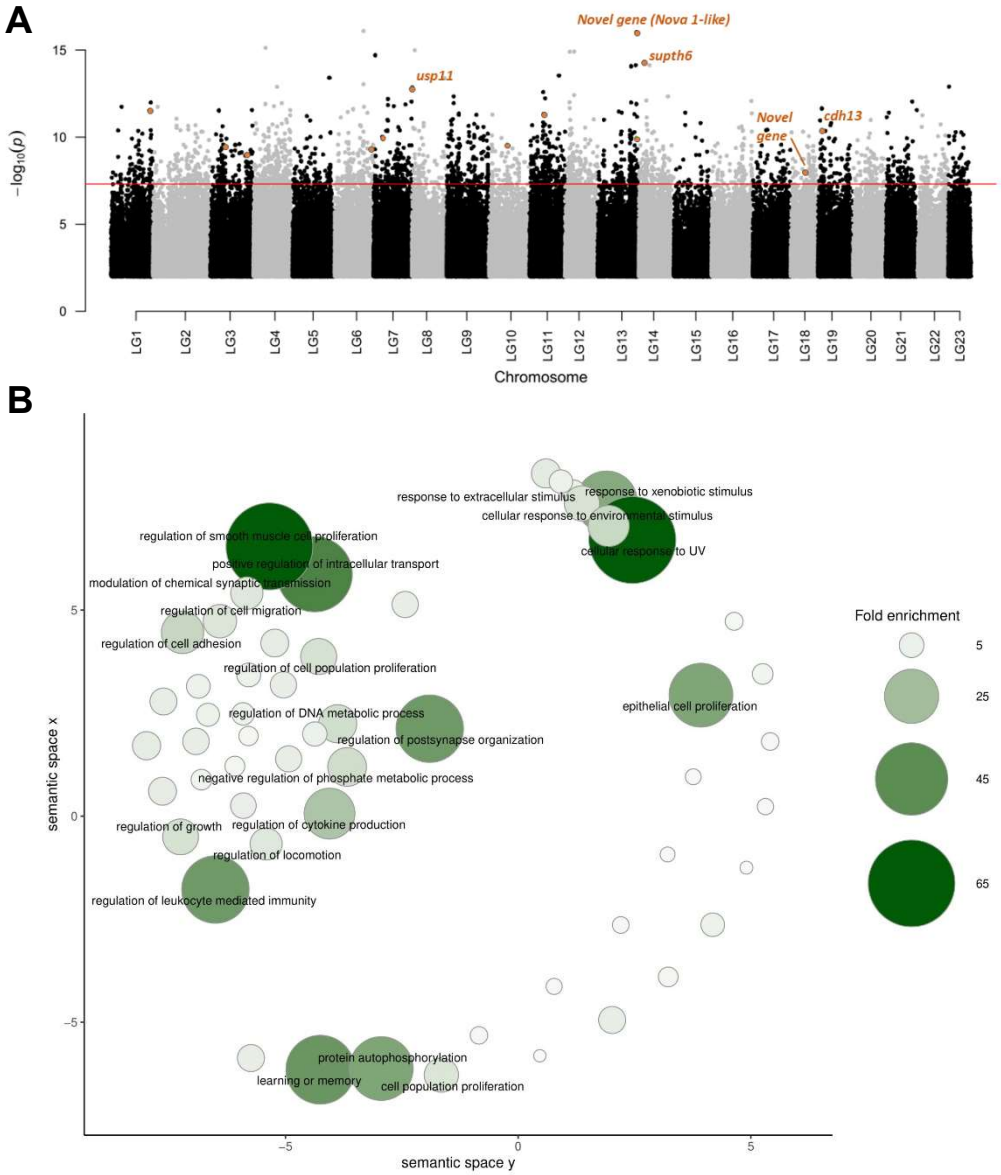

Figure 2. Genetic basis of sociability in the guppy. (A) Manhattan plot of -log10(p) values across the guppy genome resulting from a Fisher's exact test comparing allele frequency differences between high and low sociability female guppies. We merged pooled-DNA sequences of three independent replicates 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 to 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 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[74] (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[75].

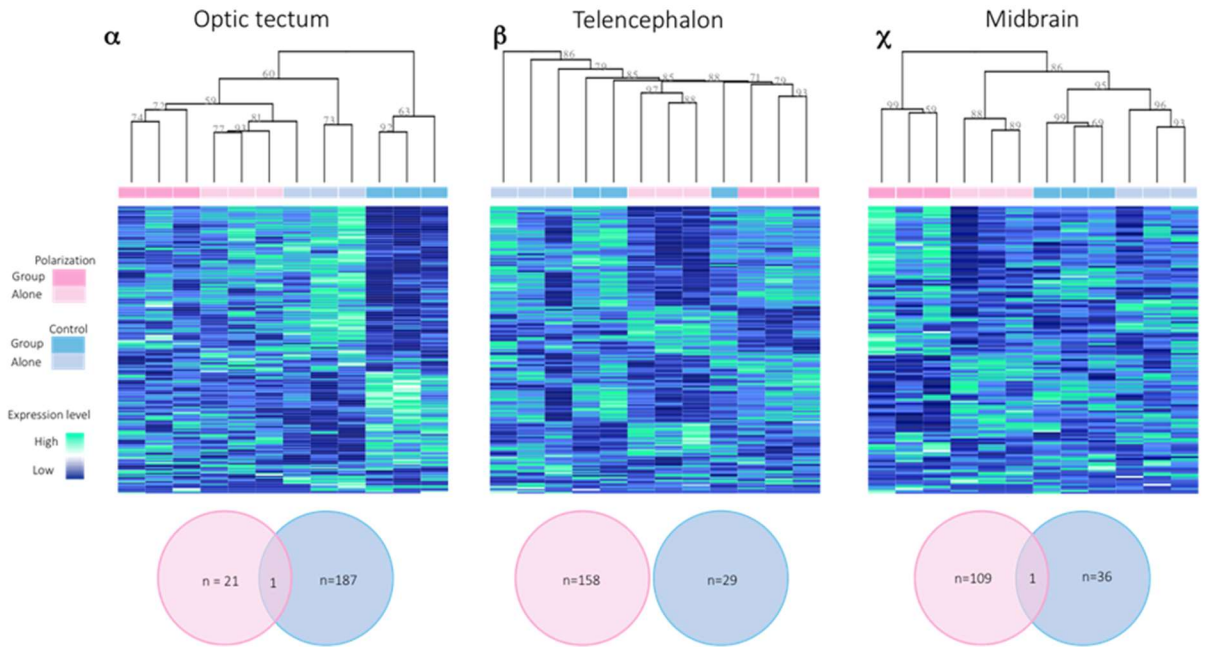

Figure 3. Neurogenomic response of schooling in guppies. Hierarchical clustering and relative expression level for all differentially expressed genes between Alone and Group treatments in the (α) optic tectum, (β) the telencephalon and (χ) the midbrain. Differentially expressed genes were identified 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 total number of differentially expressed genes in each tissue for polarization and control lines.

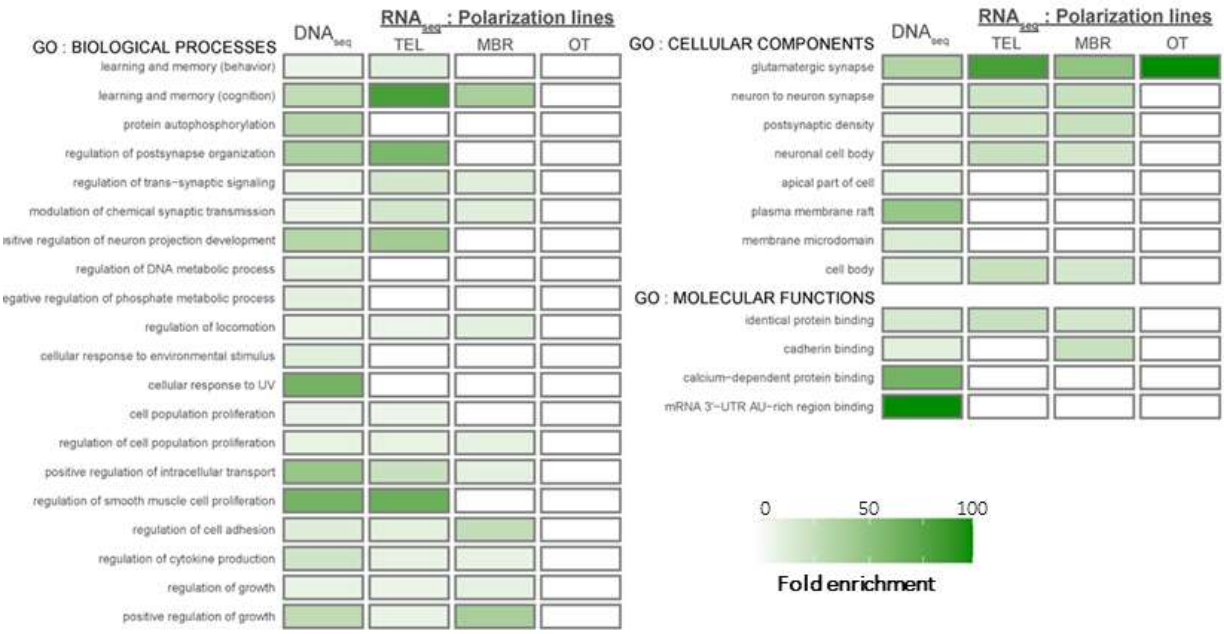

Figure 4. Functional characterization of genes of interest across experiments. Visualization of functional overlap based on Gene Ontology annotations between genes of interest highlighted in our genomic analyses of DNA comparing Pool-seq of high and low sociability female guppies (central column), and in our differential expression analyses evaluating the transcriptome of multiple brain regions of polarization-selected (right columns) and control (left columns) lines of female guppies exposed to two different social contexts, swimming alone or with a group of conspecifics. 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, Mirjam Amcoff, Severine D. Buechel, Maksym Romenskyy, Niclas Kolm, Judith E. Mank

*corresponding author: corral@zoology.ubc.ca

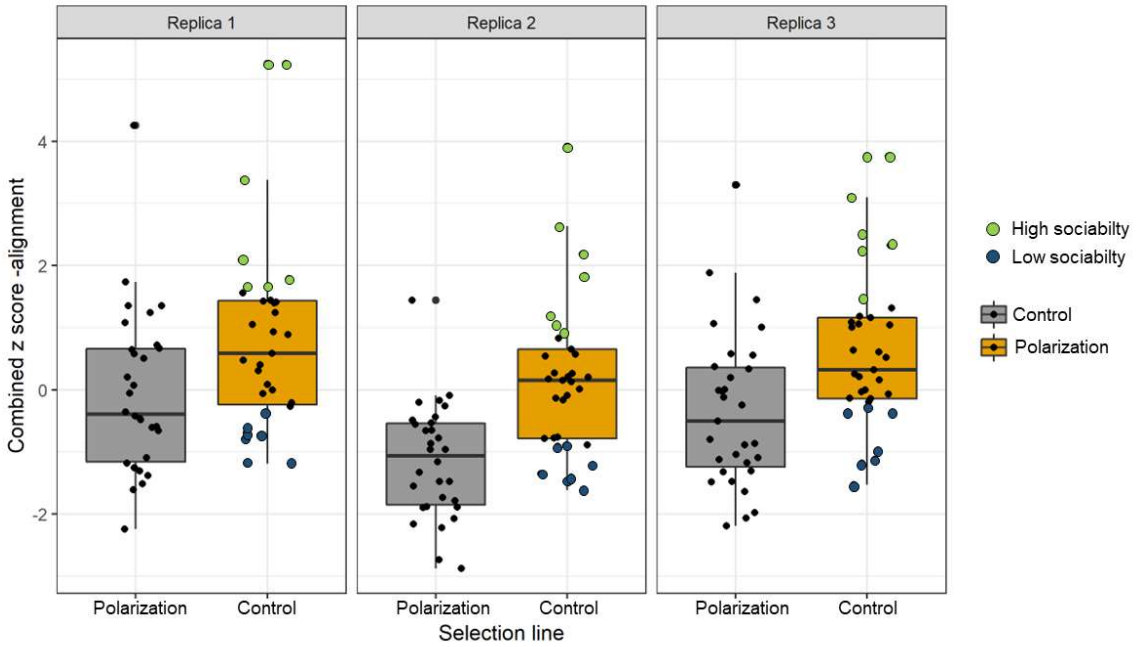

Figure S1. Mean score for alignment to group direction in females for each of the 195 families measured when evaluating the heritability of sociability by means of open field tests in which guppies from three independently replicated polarization-selected (orange) and control (gray) lines were exposed 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.

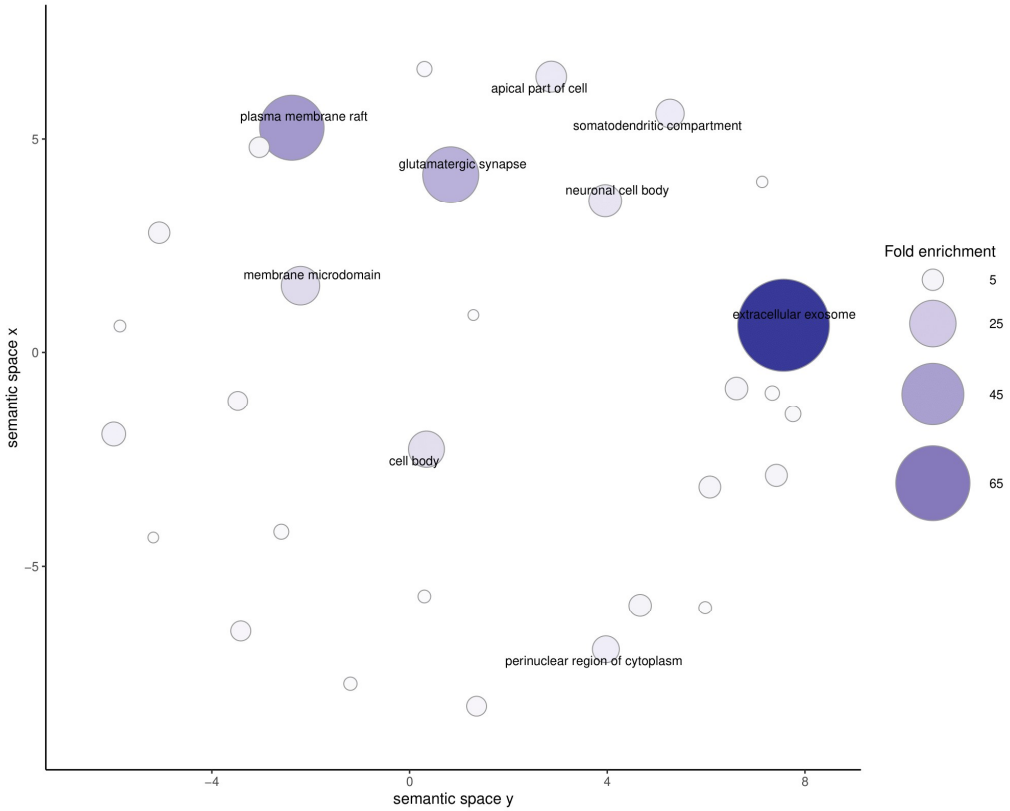

Figure S2. Clustering of statistically significant overrepresented Gene Ontology annotations for cellular components associated to differences between high and low sociability in guppies. Point size and color provide information on fold enrichment value from the statistical overrepresentation test performed 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].

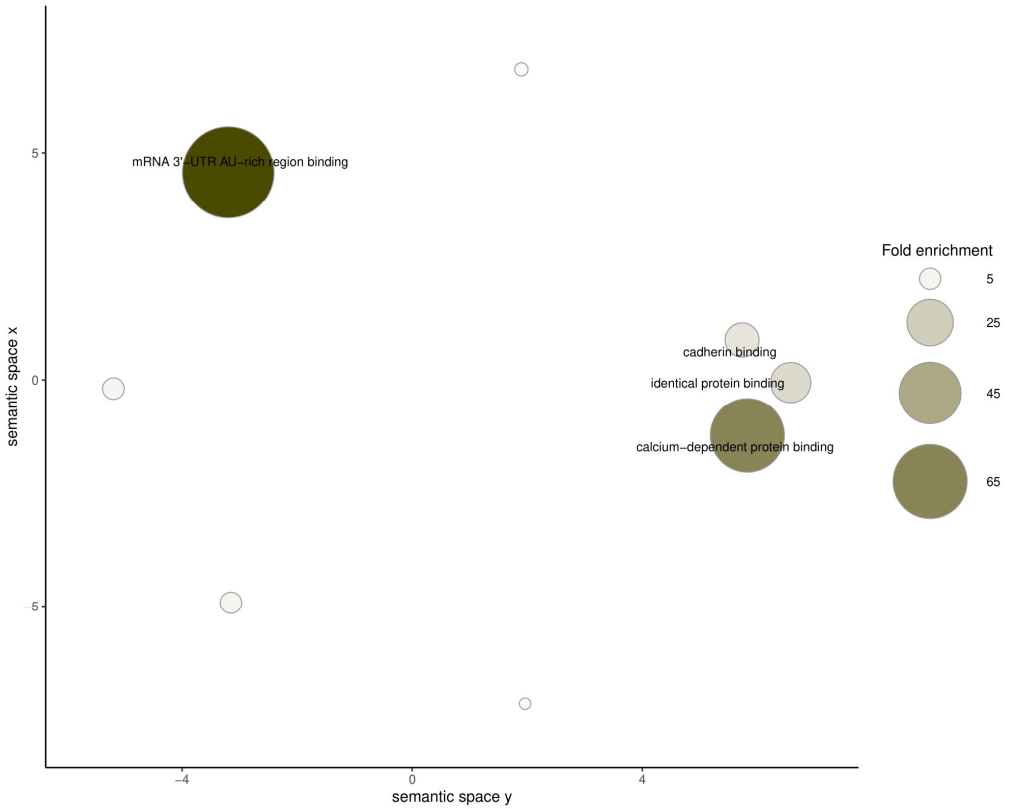

Figure S3. Clustering of statistically significant overrepresented Gene Ontology annotations for molecular functions associated to differences between high and low sociability in guppies. Point size and color provide information on fold enrichment value from the statistical overrepresentation test performed 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].

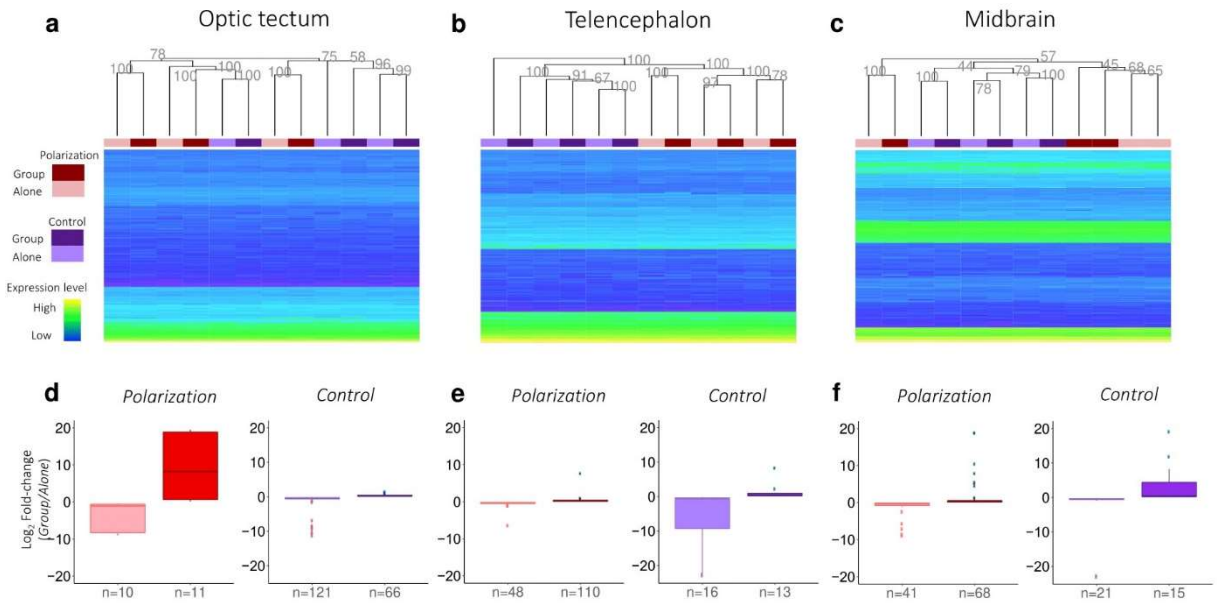

Figure S4. Top panels: Hierarchical clustering of gene expression across all significantly expressed genes the optic tectum (a), telencephalon (b) and midbrain (c) for polarization-selected selection and control lines. Values on top of nodes correspond to Approximately Unbiased bootstrap values. Heatmap 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.

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, Mirjam Amcoff, Severine D. Buechel, Maksym Romenskyy, Niclas Kolm, Judith E. Mank

*corresponding author: corral@zoology.ubc.ca

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

R syntax for statistical models

- Alignment_model = lmer(alignment ~ selection_line*sex + generation + mean_alignmnet_group + mean_speed_group + (1|selection_line:replicate) + (1 |testing order_group_) + (1 |sex:family_id)

- Attraction_model = lmer(log(nearest_neighbour_distance) ~ selection_line*sex + generation + mean_alignmnet_group + mean_speed_group + (1|selection_line:replicate) + (1 |testing order_group_) + (1 |sex:family_id)

Table S3. 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 S4. 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

R syntax for statistical models

- Alignment full pedigree model

```
#Prior:
pr_aln_model <- c(
  prior(normal(0, 3), class = 'b'),          # normal priors sd = 3 on fixed effects
  prior(student_t(5, 0, 1), class = 'sd'),   # student t priors, with df = 5 and sd = 1 on random
effects
  prior(student_t(3, 0, 3), class = 'sigma'),# student t priors, with df = 3 and sd = 3 on random
effects
  prior(lkj(1), class = 'cor')               # flat prior between -1 and 1 for rFM
)

#Animal model:

aln_model <- brm(
  align_std ~ 1 + sex * line + sex:generation + speed_std +
    (1 | group) + (1 | line:rep) + (0 + sex | animal),
  data = Data,
  prior = pr_aln_model,
  cov_ranef = list(animal = A),
  chains = 24, cores = 24, control = list(adapt_delta = 0.95), warmup = 1500, iter = 2500,
  refresh = 10)

#Heritability calculations:

    male_h2 = sd_animal__sexm^2 /
      (sd_animal__sexm^2 + sd_group__Intercept^2 + `sd_line:rep__Intercept`^2 + sigma^2),
    female_h2 = sd_animal__sexf^2 /
      (sd_animal__sexf^2 + sd_group__Intercept^2 + `sd_line:rep__Intercept`^2 + sigma^2)
```

Attraction full pedigree model

```
#Prior:
pr_att_model <- c(
  prior(normal(0, 3), class = 'b'),          # normal priors sd = 3 on fixed effects
  prior(student_t(5, 0, 1), class = 'sd'),   # student t priors, with df = 5 and sd = 1 on random
effects
  prior(student_t(3, 0, 3), class = 'sigma'),# student t priors, with df = 3 and sd = 3 on random
effects
  prior(lkj(1), class = 'cor')               # flat prior between -1 and 1 for rFM
)

#Animal model:

att_model <- brm(
  att_std ~ 1 + sex * line + sex:generation + speed_std +
    (1 | group) + (1 | line:rep) + (0 + sex | animal),
  data = Data,
  prior = pr_att_model,
  cov_ranef = list(animal = A),
  chains = 24, cores = 24, control = list(adapt_delta = 0.95), warmup = 1500, iter = 2500,
  refresh = 10)

#Heritability calculations:

    male_h2 = sd_animal__sexm^2 /
      (sd_animal__sexm^2 + sd_group__Intercept^2 + `sd_line:rep__Intercept`^2 + sigma^2),
    female_h2 = sd_animal__sexf^2 /
      (sd_animal__sexf^2 + sd_group__Intercept^2 + `sd_line:rep__Intercept`^2 + sigma^2)
```

Alignment same-sex pedigree model

```
#Prior:
pr_aln_model <- c(
  prior(normal(0, 3), class = 'b'),          # normal priors sd = 3 on fixed effects
  prior(student_t(5, 0, 1), class = 'sd'),   # student t priors, with df = 5 and sd = 1 on random
effects
  prior(student_t(3, 0, 3), class = 'sigma') # student t priors, with df = 3 and sd = 3 on random
effects)

#Animal model:
aln_samesexfemale/ aln_samesexmale/  <- brm(
  align_std ~ 1 + line + generation + speed_std +
    (1 | group) + (1 | line:rep) + (1 | animal),
  data = Data_female/Data_male,
```

```
  prior = pr_aln_model,
  cov_ranef = list(animal = A),
  chains = 16, cores = 16, control = list(adapt_delta = 0.99), warmup = 2000, iter = 4000,
 refresh = 100)
```

#Heritability calculations:

```
    female_h2 = sd_animal__Intercept^2 /
      (sd_animal__Intercept^2 + sd_group__Intercept^2 + `sd_line:rep__Intercept`^2 +
sigma^2)
    male_h2 = sd_animal__Intercept^2 /
      (sd_animal__Intercept^2 + sd_group__Intercept^2 + `sd_line:rep__Intercept`^2 +
sigma^2)
```

Alignment same-sex pedigree model

#Prior:

```
pr_att_model <- c(
  prior(normal(0, 3), class = 'b'),          # normal priors sd = 3 on fixed effects
  prior(student_t(5, 0, 1), class = 'sd'),   # student t priors, with df = 5 and sd = 1 on random
effects
  prior(student_t(3, 0, 3), class = 'sigma') # student t priors, with df = 3 and sd = 3 on random
effects)
```

#Animal model:

```
att_samesexfemale/ aln_samesexmale/  <- brm(
  att_std ~ 1 + line + generation + speed_std +
    (1 | group) + (1 | line:rep) + (1 | animal),
  data = Data_female/Data_male,
  prior = pr_att_model,
  cov_ranef = list(animal = A),
  chains = 16, cores = 16, control = list(adapt_delta = 0.99), warmup = 2000, iter = 4000,
    refresh = 100)
```

#Heritability calculations:

```
    female_h2 = sd_animal__Intercept^2 /
      (sd_animal__Intercept^2 + sd_group__Intercept^2 + `sd_line:rep__Intercept`^2 +
sigma^2)
    male_h2 = sd_animal__Intercept^2 /
      (sd_animal__Intercept^2 + sd_group__Intercept^2 + `sd_line:rep__Intercept`^2 +
sigma^2)
```

Table S5. 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