nature ecology & evolution

ARTICLES

https://doi.org/10.1038/s41559-018-0682-4

Early neurogenomic response associated with variation in guppy female mate preference

Natasha I. Bloch[1*], Alberto Corral-López[2], Séverine D. Buechel[2], Alexander Kotrschal[2], Niclas Kolm[2,4] and Judith E. Mank[1,3,4]

Understanding the evolution of mate choice requires dissecting the mechanisms of female preference, particularly how these differ among social contexts and preference phenotypes. Here, we studied the female neurogenomic response after only 10 min of mate exposure in both a sensory component (optic tectum) and a decision-making component (telencephalon) of the brain. By comparing the transcriptional response between females with and without preferences for colourful males, we identified unique neurogenomic elements associated with the female preference phenotype that are not present in females without preference. A network analysis revealed different properties for this response at the sensory-processing and the decision-making levels, and we show that this response is highly centralized in the telencephalon. Furthermore, we identified an additional set of genes that vary in expression across social contexts, beyond mate evaluation. We show that transcription factors among these loci are predicted to regulate the transcriptional response of the genes we found to be associated with female preference.

Understanding the evolution of critical animal behaviours requires identifying the underlying mechanisms by which the nervous system produces these behaviours[1–5]. Many of the most extravagant behaviours in nature are related to mate choice and reproduction. Mate choice has a major effect on organismal fitness, and is therefore subject to powerful natural selection and sexual selection pressures[6–8]. The steps involved in mating and other behaviours are mediated by changes in neural activity in the brain. Like other input from the external environment to the brain, mating stimuli are translated into neural activity triggered by acute and rapid cascades of gene expression changes. These in turn cause modifications in synaptic activity and metabolic processes, or activate further transcriptional pathways[1,9,10]. We now know that coordinated changes in the expression of many genes (the neurogenomic response[11]) are the basis of behavioural states[9,10] and play a critical role in modulating the inherent plasticity that allows our brain to respond appropriately to diverse stimuli[12,13].

Studying the gene expression changes that characterize the neurogenomic state behind mating decisions is an important part of dissecting the mechanisms underlying mating preferences and mating behaviour. Previous studies primarily based on candidate genes and/or whole transcriptomes[2,3,9] have identified some key components associated with the neural processes underlying social behaviours and mate preferences[3,14–18]. Here, our goal is to build on this knowledge by characterizing the transcriptional response triggered by different mating contexts, which is key to understanding how the brain coordinates the multitude of behaviours elicited by diverse stimuli and contexts[10,19–22]. We compared the early transcriptional response in two mating contexts: after exposure to attractive and unattractive males, in females with and without female preference phenotypes. We used the Trinidadian guppy, *Poecilia reticulata*, a model for studies of sexual selection[23–25], in which female preference and male colouration co-evolve across natural populations[26–28].

Various explanations have been offered for the association between female preference and male colour in wild guppies[29–31], but recent evidence suggests that the strength of female preference could be linked to brain size and cognitive ability[32]. Through behavioural tests on selection lines for relative brain size[33], we recently showed that females from replicate small-brained lines have convergently lost their preference for colourful males compared with wild-type and large-brained females[32]. The variation we found in the female preference phenotype in these selection lines[32] mirrors the variation among natural populations[26–28], presenting a unique opportunity to study the neurogenomics of female mating decisions comparatively while controlling for genetic background[34].

Previous studies measured whole transcriptome expression changes after 30 min of mate exposure[35], when the transcriptional response is easily detectable. However, within 10 min of mate exposure, guppy females perceive and evaluate males, experience changes in receptivity, and make a decision on whether or not to mate[23]. In order to dissect the early response of the female preference neurogenomic pathway, and to understand the transcriptional basis of variation in female preference, we used RNA sequencing (RNA-seq) to compare brain gene expression in females from different selection lines after only 10 min of exposure to a colourful (attractive) male, to a dull (unattractive) male or to another female (Fig. 1a). We focused on two brain components (Fig. 1c): the optic tectum, because it is involved in the sensory processing of visual signals, and the telencephalon, because it integrates those signals to mediate complex decision making, including social and mating decisions[36–38].

Our results reveal that guppy females with clear mate preferences exhibit a distinctive brain transcriptional response following exposure to attractive males. Genes associated with this response are more connected and central in the telencephalon co-expression network, revealing differences in the female mate preference transcriptional cascade in the various components of the brain mediating mating interactions. We also identified genes that vary across different social contexts, beyond mate evaluation, and found that these genes exhibit different expression patterns across mating and

[1]Department of Genetics, Evolution and Environment, University College London, London, UK. [2]Department of Zoology/Ethology, Stockholm University, Stockholm, Sweden. [3]Department of Organismal Biology, Uppsala University, Uppsala, Sweden. [4]These authors contributed equally: Niclas Kolm, Judith E. Mank. *e-mail: n.bloch@ucl.ac.uk

social encounters. Our results uncover the early components and structure of the genetic networks underlying female mate preference. These findings have important implications, as they provide a foundation to understand the genetics and evolution of mating decisions and mate choice.

Results

Identifying a transcriptional response uniquely associated with female preference. We first determined whether there was a transcriptional response uniquely associated with female preference. For this, we focused on those genes with significant and concordant differences in expression (DE) between attractive and dull male treatments in Preference females (that is, females with clear preferences, from wild-type and large-brained lines[32]), designated as 'X' in Fig. 1b and Supplementary Fig. 1 (see Methods for details). In order to identify genes associated with the evaluation of an attractive male, that fits the females' intrinsic preference, we further filtered these DE genes keeping only those that were also differentially expressed between attractive and female treatments, but not between dull and female treatments (area 'P' in Fig. 1b).

The resulting genes, which are associated with the female preference phenotype in Preference lines, comprised 193 genes in the optic tectum and 106 in the telencephalon (referred to as Preference DE genes; Table 1, Supplementary Table 1, Supplementary Data 1 and 2). Only eight genes were differentially expressed in both tissues. This low overlap is not surprising considering the demonstrated differences in the expression of activity-regulated genes across brain regions in birds and fish[10,39]. Even though evolutionary models predict a sex linkage of female preference genes under the good genes model[40,41], we did not observe an enrichment of these Preference DE genes on the X chromosome (LG12, $P > 0.05$). Instead, we observed enrichment of optic tectum and telencephalon Preference DE genes on various autosomes (Supplementary Table 2). As a species with Y-linked male displays, guppies may be an exception to good genes models[41]. Importantly, strong female preferences could also evolve from direct selection on the sensory system[42] or, as we hypothesize, on cognitive ability[32].

Preference DE genes exhibited a distinct transcriptional signature in Preference females exposed to an attractive male in both tissues, and thus clustered together separately from all the other samples (Fig. 2). However, it is important to note that in the optic tectum, Non-preference samples showed differences in the expression of Preference DE genes, similar to those seen in Preference females exposed to a dull male or a female (Fig. 2). There is therefore some activity for Preference DE genes in Non-preference females at the sensory-processing level, suggesting that the difference in attractiveness between the two male types is being perceived and processed by Non-preference females. We did not observe this pattern at the decision-making level, in the telencephalon. Here, Non-preference samples grouped in a third separate cluster, where Preference DE genes did not show any differences in expression. This suggests that Preference DE genes in the telencephalon are not recruited to the decision-making process in Non-preference females. We know these differences are due to the social stimuli, as samples do not follow the same clustering pattern when transcriptome-wide expression is considered (Supplementary Fig. 2). Moreover, we have previously characterized the genetic differences between large-brained (Preference) and small-brained (Non-preference) lines, and showed that they only differ in the regulation of one locus, Angiopoeitin-1(*Ang-1*)[34]. Expression of this key gene during development influences both the relative brain size and neural density of these fish. We suggest that this developmental difference is indeed the main driver of the variation in brain size among selection lines[34].

We next performed an identical differential expression analysis and filtering in Non-preference females. We found only 61 and 38 loci that were differentially expressed between the attractive and dull male treatments in the optic tectum and telencephalon, respectively (Non-preference DE genes; Table 1, Supplementary Data 1 and 2). Although members of the same gene families were differentially expressed in lines with opposing preference phenotypes (sodium calcium exchanger proteins and ribosomal proteins, among others; Supplementary Table 1), none of these overlapped with Preference DE genes. Unlike Preference DE genes, Non-preference DE genes did not exhibit a distinct expression signature in Preference females (Supplementary Fig. 3) and were enriched in different chromosomes to those of Preference DE genes (Supplementary Table 2).

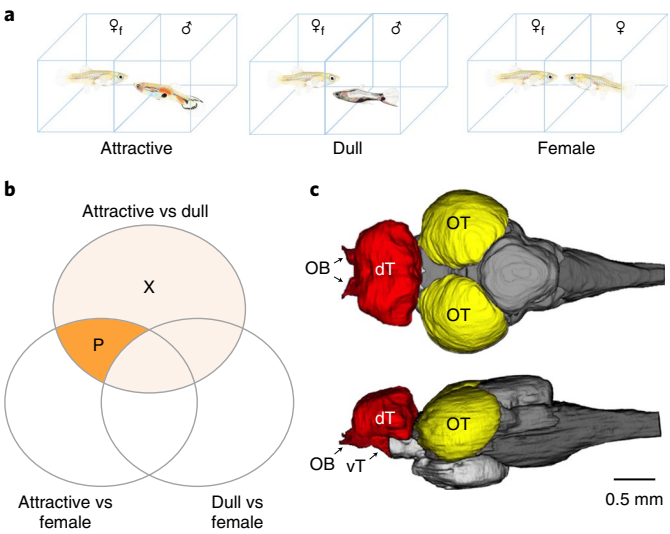

Fig. 1 | Experimental set-up used to find neurogenomic pathways associated with mate preferences. a, Diagram of the three treatments. Focal females (♀f) were exposed to an attractive male (left), a dull male (centre) or another female as a control condition (right). Note, guppies are not drawn to scale. **b**, Venn diagram illustrating the various pairwise comparisons used to identify differentially expressed genes between treatments. Identification of differentially expressed genes and permutations were performed for each pairwise treatment comparison and separately for Preference and Non-preference lines in both tissues. See Table 1 for the results of all comparisons. Area X indicates all the genes differentially expressed between the attractive and dull treatments, and P is the final set of Preference DE genes after filtering to keep only those attractive versus dull DE genes that are also differentially expressed in the attractive versus female comparison but not in the dull versus female (see Methods for details). **c**, Schematic representation of a top view (top) and lateral view (bottom) of the major regions of the guppy brain. We examined gene expression in the optic tectum (OT) and the telencephalon, which included the dorsal telencephalon (dT), the ventral telencephalon (vT), the preoptic area and olfactory bulbs (OBs). The latter are less than 2.9% of the mass. The optic tectum samples included the laminated superior area of both hemispheres.

Co-expression network attributes and modularity of the female preference neurogenomic response. We next investigated gene relationships in the context of weighted co-expression networks (WGCNA)[43,44] for each tissue separately. Co-expression networks allow us to examine the regulatory connections between differentially expressed genes and to determine the modular structure of transcriptional responses[45]. The optic tectum and telencephalon networks retained 6,297 genes and 3,540 genes, respectively (Supplementary Table 3; Supplementary Fig. 4; see Methods). For subsequent analyses, we focused on DE genes remaining in the co-expression networks, as these genes have strong transcriptional connections, a characteristic we might expect for genes at the apex

Table 1 | Differentially expressed genes

		Attractive vs dull	Attractive vs female	Dull vs female	Total (unique genes)
Optic tectum					
Preference	Attractive vs dull genes that pass the permutation 5% threshold	1,278 (X)	1,125	982	2,746
	Preference DE genes (after filtering[a])	**193 (P)**	–	–	–
Social DE genes					**357**
Non-preference	Genes that pass the permutation 5% threshold	842 (X)	1,973	1,449	3,393
	Non-preference DE genes (after filtering[a])	**61 (P)**	–	–	–
Telencephalon					
Preference	Genes that pass the permutation 5% threshold	919 (X)	746	785	1,999
	Preference DE genes (after filtering[a])	**106 (P)**	–	–	–
Social DE genes					**161**
Non-preference	Genes that pass the permutation 5% threshold	847 (X)	705	677	1,853
	Non-preference DE genes (after filtering[a])	**38 (P)**	–	–	–

Letters in parentheses refer to Venn diagram sections highlighted in Fig. 1. [a]Genes that were considered differentially expressed between attractive and dull treatments following the permutation 5% cut-off were filtered for concordant expression across all the replicate lines and for differential expression between attractive versus female and dull versus female keeping only genes in section P of Fig. 1. See main text for further details. Values in bold correspond to the final set of filtered Preference DE genes, Non-preference DE genes and Social DE genes.

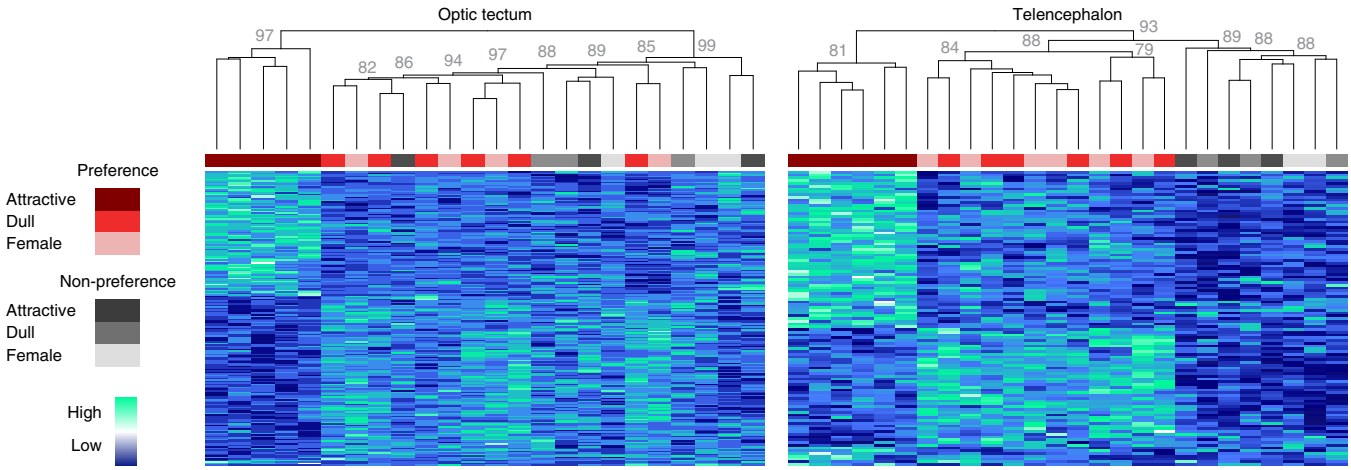

Fig. 2 | Hierarchical gene-expression clustering of Preference DE genes. Hierarchical gene-expression clustering of samples for Preference DE genes differentially expressed between attractive and dull male treatments in the optic tectum ($n=193$) and telencephalon ($n=106$). Colours below the dendrogram correspond to the sample treatment and line as outlined in the key. Values on the top of the nodes correspond to bootstrap Approximately Unbiased *P* values, computed by multiscale bootstrap resampling[91] (all bootstrap values >70%, those <80% not shown for clarity).

of genetic pathways involved in the female preference response. In order to investigate the network properties of DE genes relative to genes with known roles in social behaviour, we compiled a list of genes previously shown to have roles in social or mating behaviour and mate preferences (Supplementary Table 4), including synaptic plasticity genes (SPGs), some of which are immediate early genes (IEGs) (Supplementary Table 5). The context- and stimulus-dependent plasticity that characterizes the brain, allowing it respond differently to thousands of stimuli, is due in part to the response of these genes which alter synaptic connections[12,18,46,47].

We found that Preference DE genes in the optic tectum and in the telencephalon networks have different properties. Our analysis of network attributes revealed that Preference DE genes in the optic tectum are distributed throughout the co-expression network with highly variable centrality and connectivity measures (Table 2). In contrast, Preference DE genes are both centrally and highly connected in the telencephalon network (Table 2; Supplementary Fig. 4). This suggests that the evaluation of males of different qualities causes responses with different characteristics at the sensory-processing and the decision-making levels. The greater centrality and connectivity of Preference DE genes in the telencephalon suggests that we have identified upstream control genes in the decision-making component of the brain, genes that are responsible for initiating the transcriptional cascades underlying female preference behaviours. These ultimately lead to the decision to mate, downstream endocrine responses and changes in future behaviour. Crucially, this pattern was not observed in the telencephalon of Non-preference females in response to an attractive male.

We also found that genes previously associated with mate preference and social and mating behaviour[3] (Supplementary Table 4) were significantly more peripheral (that is, genes with lower gene connectivity at the periphery of the co-expression network) than our Preference DE genes in the telencephalon (Supplementary Fig. 4). This finding is consistent with the notion that the telencephalon Preference DE genes we identified after 10 min of treatment exposure are the upstream components of the preference pathway and

Table 2 | Co-expression network centrality and connectivity measures

a, Comparison of centrality and connectivity measures between the optic tectum and telencephalon networks		*n*	Optic tectum	*n*	Telencephalon	*P* value
Preference DE genes	Degree average[a]	57	3.56 (2.83)	12	8.67 (3.64)	0.02*
	Clustering coefficient[b]		0.16 (0.72)		0.53 (0.53)	<0.001**
	Neighbourhood connectivity[c]		7.84 (3.30)		21 (3.66)	<0.001**
Non-preference DE genes	Degree average[a]	31	6.48 (3.17)	6	3.83 (2.10)	NS
	Clustering coefficient[b]		0.24 (0.56)		0.49 (0.70)	NS
	Neighbourhood connectivity[c]		11.12 (3.5)		8.89 (3.56)	NS
Social affiliation/female preference genes	Degree average[a]	10	13.8 (3.7)	3	1.7 (0.4)	0.02*
	Clustering coefficient[b]		0.34 (0.6)		0 (0)	<0.01**
	Neighbourhood connectivity[c]		21.3 (3.7)		2.5 (0.5)	<0.01**
b, Comparison of centrality and connectivity measures between Preference DE genes and Social affiliation/female preference genes		**Optic tectum**			**Telencephalon**	
Social affiliation/female preference genes compared to Preference DE genes	Sample sizes	57/10			12/3	
	Degree average[a]	$P=0.04$*			$P=0.02$*	
	Clustering coefficient[b]	$P<0.01$**			$P<0.001$**	
	Neighbourhood connectivity[c]	$P<0.01$**			$P=0.02$*	

All *P* values correspond to *t*-tests. Sample sizes in b correspond to Preference DE genes/social affiliation and female preference genes. Numbers in parentheses indicate standard deviations. NS, not significant. [a]The number of edges, that is other genes, each gene is connected to within the network. Central genes in the network will therefore have high degree values as opposed to more peripheral network genes. [b]The ratio of the number of edges between the neighbours of a gene, and the maximum number of edges that could possibly exist between such neighbours (number between 0 and 1). This is a measure of how connected a gene is relative to how connected it could be given the number of neighbours it has. This value will approach 0 for a loosely connected gene and 1 for a fully connected gene in the centre of a network. [c]The average connectivity across all neighbours.

could induce the expression of genes that have been identified by previous work focused on 30 min of treatment exposure.

We next identified gene modules in our co-expression network, which represent clusters of genes with highly correlated expression[44,48] (Supplementary Table 3; Supplementary Fig. 5). Co-expression network modules are a powerful tool in this context, as genes within the same module have been experimentally shown to share functions and/or biological processes[45,49]. In the optic tectum, five modules (modules OT9, OT12, OT15, OT21 and OT24; Fig. 3a) are enriched in Preference DE genes, and three separate modules are enriched in Non-preference DE genes (OT16, OT31, OT32; Fig. 3a) . See Supplementary Table 6 for Gene Ontology (GO) terms associated with these modules.

Module OT24 is particularly interesting, as it is enriched in Preference DE genes that show strong transcriptional connections to multiple genes known for their role in female preferences in this module and in module OT17. Preference DE genes in this subnetwork include *gria3*, a member of the AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid) glutamate receptor family known to be an important component of the female preference response[50]. Other Preference DE genes include the following: *scn2a* and *scn8a*, which are known to have molecular functions in brain circuits that mediate specific behaviours[51]; *agap3*, which is involved in signal transduction; *syn1*, which is known to be involved in synaptic plasticity and social behaviour[52]; *baz2a*, which regulates transcription of androgen receptors; and *slc24a2*, a critical gene in signal transduction[53] with known roles in cognition and memory[54], and a target of the IEG *fosl1*. The network structure revealed that these genes are connected to other known components of the female preference transcriptional response[3,18], including *neuroligin-2*, *neuroligin-3*, *stmn2a* and *stmn2b*. Such connections, in conjunction with the elevated connectivity and centrality scores, suggest that the Preference DE genes we identified may act to coordinate the transcriptional response behind female preferences documented in previous studies, thus supporting their roles in the initiation of neural and behavioural cascades of female mating decisions.

Once the visual signal travels from the optic tectum into the telencephalon, we see further separation of modules grouping Preference DE genes and modules associated with Non-preference DE genes. In the telencephalon, modules T4, T37 and T46 are significantly enriched in Preference DE genes, while modules T23, T29 and T31 are enriched in Non-preference DE genes (Fig. 3b). Although not enriched in Preferences DE genes, module T13 is worth noting as it connects three Preference DE genes (out of 12 in total) with a very large number of SPGs and IEGs (Fig. 3b). Among the modules enriched in SPGs and IEGs and social behaviour/female preference genes (T2, T12, T13, T32 and T43), modules T12 and T43 group SPGs and IEGs and genes identified as regulators of female preferences at 30 min[15,55] that could be activated downstream of the Preference DE genes we identified.

Function and regulation of differentially expressed genes. We found that genes in modules associated with the neurogenomic response of female preference are enriched in pathways underlying neural plasticity[13], including Ras signalling and long-term potentiation pathways, the Wnt signalling pathway, the neurotrophin signalling pathway and phototransduction (Supplementary Table 7). Module OT24 in particular is enriched in GO terms highly relevant to behaviour, memory and learning, including the glutamate receptor signalling pathway (Supplementary Table 6). We also found that different optic tectum modules are regulated by different sets of transcription factors (TFs), and that many of the Preference DE genes are predicted to have TF motifs for the IEGs *egr1*, *egr2, c-fos* and *c-jun*, as well as the neuronal plasticity and long-term memory modulator *CREB* (Supplementary Fig. 8).

Telencephalon Preference DE genes include several ribosomal proteins and genes involved in hormone signalling and response, such as *eef2* and *c2cd5* (Supplementary Table 6). A promoter analysis showed enrichment for TF motifs for *CREB* and *srf*, both part of the Ca^{2+}/calmodulin-dependent kinase (CaMK) signalling pathway and central regulators of neural plasticity and memory[56], as well as *pitx2*, among others (Supplementary Fig. 8). Aside from

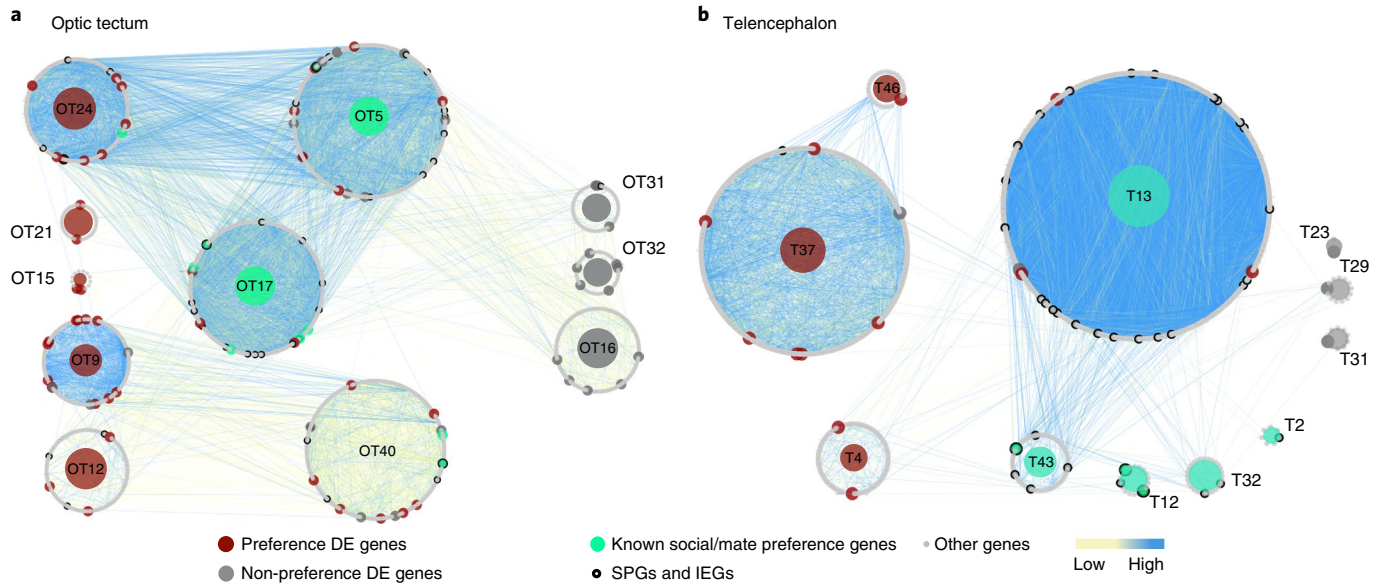

Fig. 3 | Optic tectum and telencephalon co-expression networks' module overview. a, Optic tectum. **b**, Telencephalon. Each circle of genes represents a module and the dots forming the module circle represent genes. The size of each module is therefore proportional to the number of genes in that module. The colour of each dot refers to its DE category or functional affiliation as shown in the key. Numbered modules are referred to in text, and correspond to modules after merging (Supplementary Fig. 5). Modules significantly enriched for Preference DE genes are highlighted in red for Preference lines and grey for Non-preference lines. Modules highlighted in green are significantly enriched in known social behaviour/mate preference genes and/or SPGs. Edge connections are highlighted according to weight, with stronger connections, for correlations approaching 1 or −1, shown in blue. Modules with no differentially expressed genes or behavioural genes of interest, as well as edges associated with these modules, are hidden for clarity.

ribosomal proteins, all the genes had TF motifs for the IEGs *c-fos* and *c-jun*, TFs previously associated with activity levels in brain regions mediating various behaviours, including social interactions (Supplementary Fig. 8).

Preference DE genes in modules OT17 (*npr2*) and T37 (*eef2*) have roles in downstream hormone secretion and signalling, being located upstream within the oxytocin signalling pathway, as well as genes in module OT21 (*tubb4a* and *tmem198*) in the gonadotropin-releasing hormone (GnRH) receptor pathway, shown to have an important role shaping preferences during interactions with potential mates[57,58] (Supplementary Table 1). These genes could be responsible for controlling the female physiological changes associated with preparation for mating and reproduction.

Identifying genes that vary in expression in different social interactions. In order to identify genes modulating social interactions beyond mate evaluation, we determined which genes were differentially expressed across all social interactions in all females, independent of their preference phenotype (in Preference and Non-preference lines; Supplementary Fig. 1). We found 357 such DE genes (denoted Social DE genes) in the optic tectum and 161 in the telencephalon (Table 1; Supplementary Fig. 6).

We examined overall differences in the expression patterns of Social DE genes across treatments and lines using principal component analysis (PCA). We found that in both tissues, Preference females exposed to an attractive male exhibited a unique transcriptional signature and clustered as a separate group from the rest of the sample groups based on the first three principal components (Fig. 4). Beyond this, the pattern is different in both tissues. In the optic tectum, except for the attractive treatment in both Preference and Non-preference females, the expression of Social DE genes in different treatment groups was mostly overlapping (Fig. 4a,b). Unlike the optic tectum, principal component 1 (PC1) in the telencephalon initially separated samples by preference phenotype (Fig. 4c). PC2 and PC3 revealed a unique transcriptional pattern in Preference females exposed to an attractive male. Non-preference females lacked this unique response to attractive males, so that all male treatments clustered together (Fig. 4d). This suggests that exposure to an attractive male does not trigger a distinct transcriptional response in the telencephalon of Non-preference females.

Social DE genes include genes related to synaptic plasticity, learning, memory and social behaviour, such as *grin1* (a critical subunit of N-methyl-D-aspartate receptors (NMDAR)), *bdnf*, *neurod2*, *c-fos* and *egr2b*[13,16,18,50,59–61]. Social DE genes in both tissues are linked in several pathways relevant in behaviour, such as the Ras signalling pathway, the Wnt signalling pathway, the gonadotropin-releasing hormone (GnRH) receptor pathway and the corticotropin-releasing factor receptor signalling pathway, among others (Supplementary Table 8). A promoter region analysis[62] suggested that Preference DE genes in the optic tectum and telencephalon co-expression networks have TF motifs for our Social DE genes (Supplementary Table 9), indicating that differences in the expression of Social DE genes may trigger distinct transcriptional cascades in the different mating and social contexts of our experiment (Supplementary Fig. 7; Supplementary Table 9).

Discussion

Our goal was to characterize the neurogenomic response of female preference by identifying the differences in gene expression triggered by different mating contexts in females with and without a preference for colourful males[32]. This comparative framework allowed us to investigate which elements of the response differ in females that lack preference for attractive males[32], thus identifying the neurogenomic basis of variation in female preferences that are key to sexual selection and sexual conflict. We specifically targeted genes involved in the early female preference neuromolecular response by studying the transcriptional changes after only 10 min of mate exposure.

In both the optic tectum and telencephalon, we identified genes that differ in expression in different social contexts (Fig. 4) and

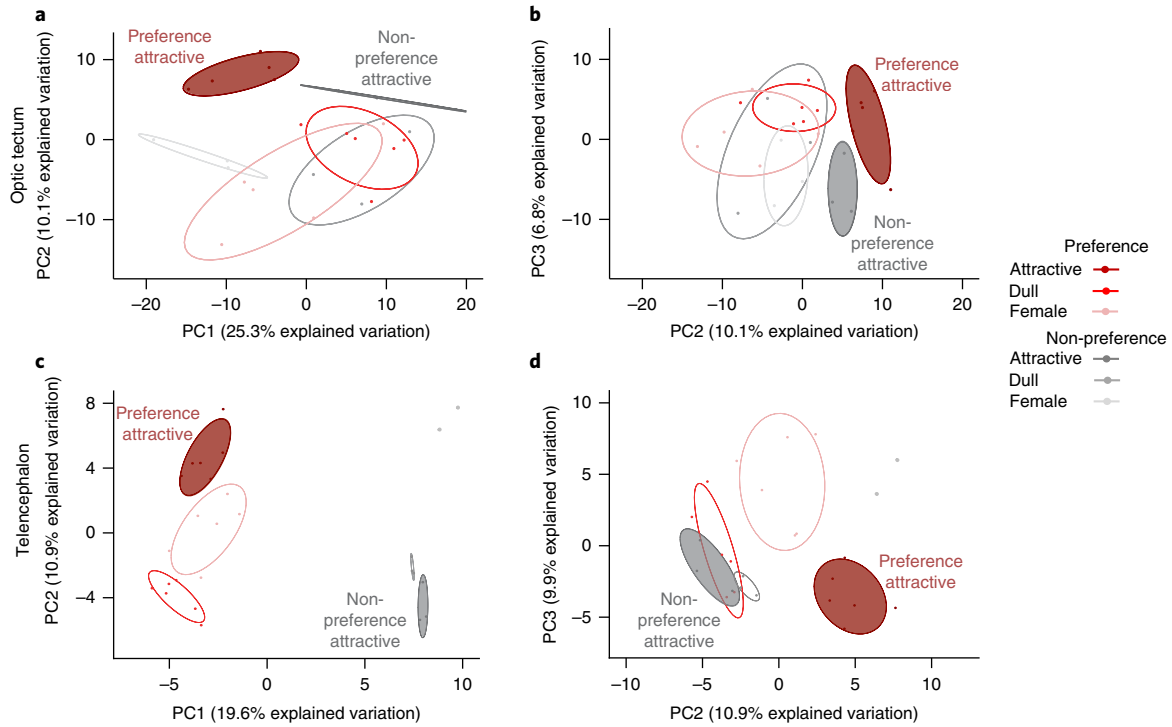

Fig. 4 | Differential transcriptional signature of Social DE genes in females exposed to attractive males. a–d, Principal component analyses of Social DE genes in the optic tectum (**a**,**b**; $n = 347$) and the telencephalon (**c**,**d**; $n = 161$). Points represent samples for each treatment/line group. In **a** and **c**, the two first principal components are plotted, and in **b** and **d**, PC2 is plotted against PC3, with the proportion of variance explained by each component printed next to the axes labels.

found evidence that the TFs among these genes probably act as neuromolecular switches triggering distinct neurogenomic states that form the basis of mating decisions and social behaviours. Consistent with this idea, we found multiple genes with unique transcriptional signatures in Preference females exposed to an attractive male, suggesting that they are part of the neurogenomic response of female preference (Fig. 2). These Preference DE genes are assembled into discrete genetic modules in the optic tectum and telencephalon, revealing the structure of the transcriptional response uniquely associated with female preference, as well as connections to other genes known to have regulating roles in social behaviour, mate preferences, learning and memory (Fig. 3).

The centrality and connectivity of Preference DE genes in the optic tectum and telencephalon showed that the properties of the response are different in both brain tissues. While we saw a diffuse response associated with female preference at the sensory-processing level, with DE genes at all levels of the network, we observed a highly centralized response for DE genes in the decision-making telencephalon. In addition to highlighting differences in the properties of the response at the sensory-processing and decision-making levels, a highly centralized response in the telencephalon is exactly what we would expect of the genes that initiate the female preference transcriptional response leading to the alternative mating decisions that follow.

Furthermore, Preference DE genes have similar expression patterns in females with and without preferences in all but the attractive male treatment at the sensory-processing level (optic tectum), suggesting that Non-preference females do perceive differences between both types of males. However, at the decision-making level (telencephalon), Preference DE genes are not activated in response to any social interactions in Non-preference females (Fig. 2). These findings, combined with the expression pattern of Social DE genes (the PCA in Fig. 4), where we see a strong differentiation in telencephalon expression between lines with different preference phenotypes along PC1, suggest that there are crucial differences in the neurogenomic response behind social and mating behaviours in the telencephalon. The expression differences seen along PC1 at the decision-making level could be a reflection of the proven differences in cognitive ability between lines[33] and is consistent with the notion that cognition plays an important role in mating decisions[37,46].

Herbert[63] originally introduced the idea that limited genetic elements can encode the multiple behaviours required to appropriately respond to various stimuli in different social and mating contexts via a complex combination of spatial and temporal activation in different brain nuclei. Here, we see evidence for a group of genes that have different expression levels in various mating contexts grouped in several discrete modules associated with female preferences, revealing the modularity of the neurogenomic preference response we observed. We see further evidence of how the brain can flexibly respond to different stimuli in the observation that multiple SPGs and IEGs are present in our Social DE genes, including *grin1*, *march8*, *bdnf*, *thoc6*, *cant1* and *thap6* in the optic tectum and *inhba, neurod2, smarcc1, c-fos, egr2b* and *thap6* in the telencephalon. However, different social behaviours have been shown to be characterized by different patterns of gene activity across the different nodes of the telencephalon forming the social decision-making network[64,65], rather than the gene activity of a single node. It would therefore be a useful avenue for future research to continue to dissect how the brain mediates its response to mating stimuli by examining detailed patterns of expression of Preference DE genes and Social DE genes across the different nodes of the telencephalon.

The comparative framework we used here enabled the identification of genes and gene modules associated with variations in female preference, and which of these are likely to factor in the neurogenomic response behind female mate choice. These findings provide

a clear testable hypothesis to investigate the mechanisms underlying the repeated and independent evolution of divergent female preference for colourful males across wild guppy populations[23,26,66,67]. Together, our results revealed the unique transcriptional response related to the earliest stages of female preference behaviour, showed the modularity of this response and identified the potential regulatory basis of this transcriptional response. Our approach and results provide a strong comparative framework for studies of the conservation of mate preference transcriptional networks across populations and species.

Methods

Study system. Guppies used in our experiment are laboratory-raised descendants of Trinidad guppies sampled from the high predation populations of the Quare River (Trinidad). We based our study on guppies from this wild-type population and six selection lines, derived from the wild-type fish, which have been selected on the basis of relative brain size. In summary, fish were indirectly selected based on parental brain size, achieving a difference of up to 13.6% in relative brain size among three replicate lines selected to have small brains, here denoted small-brained lines, and three replicate lines selected to have larger brains (large-brained lines)[33,68]. All the details on the selection experiment have been previously published[33]. Brain size in these lines has been shown to carry significant costs and benefits, conferring better cognitive abilities and better response to predators in large-brained lines[33,68,69]. These differences, however, are not likely due to the accumulation of deleterious alleles in small-brained lines, as these were shown to be more fecund[33], to have a better immune response[70] and faster juvenile growth[71]. We recently showed that females from wild-type and selection lines have measurable differences in their female preference for colourful males. While females from large-brained lines have maintained the clear female preference for colourful males seen in the wild-type line, small-brained females lack this preference[32]. We demonstrated that this difference in preference phenotype is not due to differences in opsin sequence or expression in the retina, or to variation in colour perception across lines[32].

For this study, we used virgin females from the fifth generation of selection, all aged approximately 6 months. None of the females used in this experiment was used for other behavioural experiments before this study. Fish were raised at a water temperature of 25 °C with a 12:12 light:dark schedule, and fed an alternating daily diet of flake food and live *Artemia* (brine shrimp). After the first onset of sexual maturation, females were placed in 12-litre tanks in groups of 10 fish. All tanks contained gravel, biological filters and Java moss (*Vesicularia dubyana*). In addition, we allowed visual contact between tanks containing females to enrich the social environment, but females never saw a mature male before the experiment. Experiments were performed in accordance with ethical permits approved by the Stockholm Ethical Board (reference numbers N173/13, 223/15 and N8/17). These applications are consistent with the Institutional Animal Care and Use Committee guidelines.

Preference tests. *Selection of presentation males.* For our study, we divided females among three treatments: two treatments represented a male evaluation context, in which females were presented with either an attractive male (attractive treatment) or an unattractive male (dull treatment), and a third treatment in which females were exposed to another female, representing a general social interaction treatment. Previous studies have demonstrated that females are attracted to males with brighter and larger orange areas and longer tails[23]. Following previously described general methods[32], we selected 30 wild-type males from the laboratory population stock for their colourful or dull patterns based on visual inspection. Next, these 30 males were anaesthetized with a low dose of benzocaine and photographed on both sides using a Nikon D5300 camera. We scored total colouration, body length and tail area of each male using ImageJ software v.1.44[72]. Then, we selected the four males with highest and lowest colouration that could be matched by body length. Before the trial, we made sure that these males were sexually mature by housing them together with females that were not part of the experiment and observing their sexual behaviour. As colour patterns might change over time in young fish, we repeated the whole procedure after 5 days of the experiment. In total, we used three sets of colourful–dull males during the experiment. On average, the 12 selected colourful males presented 23% more total colouration and 16% larger tails than the 12 dull males.

Behavioural treatments. We used a total of 45 wild-type females, 45 large-brained females and 45 small-brained females divided equally across the three treatments. For the selection lines, we used five females each from the three replicates. We allowed each focal female to observe the presented fish for only 10 min before ending the experiment based on our findings in a previous female mate choice study of these lines[32]. This timeframe was chosen based on previous studies[32] as an early time point at which differences in female behaviour could be observed. This short presentation time also minimizes the possibility of habituation to the experimental set-up. Preference tests were carried out in a divided tank (84 × 40 × 20 cm), which controlled for the focal female perceiving any chemical or mechanical signals. All fish were netted and transferred to their respective experimental tanks 24 h before the start of the experiment for acclimation. We ensured that all females used in gene expression analyses showed sexual interest in the males offered. For this, all trials were followed by an observer through a live broadcast of the experimental set-up in a separate room to avoid disturbances. For consistency, all trials were conducted on 15 consecutive days. Focal females belonging to the same replicate selection line and the same treatment were presented with different males to avoid uncontrolled male-driven changes in expression. For this, we balanced the number of large-brained, small-brained and wild-type females presented to colourful males, dull males and females, respectively, per day (nine trials per day). We have previously shown that our selection lines do not significantly differ in any behaviour and movement patterns in mating contexts and/or during the preference tests[32,73,74,75]. This extensive work showed no evidence for any behavioural differences in perception, activity or swimming behaviour that could affect the results.

At the end of each trial, females were euthanized by transfer to ice water. After 45 s, and with the aid of a Leica S4E microscope, we removed the top of the skull to expose the brain. We cut the olfactory and optic nerves and extracted the following forebrain regions: dorsal telencephalon, ventral telencephalon (harbouring the preoptic area) and olfactory bulbs. 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. The thalamus region was excluded from our samples. As the olfactory bulbs are very small in guppies (typically <2.9 % of the forebrain mass[76]), we use 'telencephalon' when relating to samples extracted from these forebrain regions. Next, after detachment of the cerebellar region, we dissected out the laminated superior area of the optic tectum (Fig. 1c). Dissection procedures took place in ice water within 3 min. The telencephalon and optic tectum tissue samples were immediately preserved in RNAlater (Ambion) at room temperature for 24 h and then at −20 °C until RNA extraction.

RNA extraction and sequencing. In order to recover sufficient RNA for RNA-seq, we pooled tissue from five individuals. For consistency, samples were pooled combining tissue for the same individuals for the optic tectum and telencephalon. This produced three replicate pools per treatment for each wild-type line, large-brained line and small-brained line for the optic tectum and the telencephalon (three pools per treatment per line generating nine pools per line and thus 27 pools in total for each tissue). Each sample pool was homogenized and RNA was extracted using Qiagen's RNAeasy kits following the manufacturer's standard 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 10 lanes on an Illumina HiSeq 4000. We obtained on average 52 million 75 bp read pairs per sample (47.1 million read pairs minimum, 72 million maximum).

Assembly construction. *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 using Trimmomatic v.0.35[77]. 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. After trimming, we had a total of approximately 537.6 million trimmed read pairs, 44.8 on average per individual (minimum of 36.2 million trimmed read pairs and a maximum of 56.2 million trimmed read pairs).

De novo assembly. Because the current guppy genome annotation is incomplete[78], we constructed a de novo transcriptome assembly in order to include loci that might be missing from the current annotation. All forward and reverse reads were pooled and assembled de novo using Trinity v.2.2[79] with default parameters. We filtered the resulting assembly for noncoding RNA using medaka (*Oryzias latipes*) and Amazon molly (*Poecilia formosa*) noncoding RNA sequences as reference in a nucleotide BLAST (Blastn). After eliminating all sequence matching noncoding RNAs, we picked the best isoform for each transcript. We defined the best isoform as the one with the highest expression as estimated by mapping the reads to the de novo assembly using RSEM v.1.2.20[80]. Finally, we used Transdecoder (Transdecoder v.3.0.1; http://transdecoder.github.io) with default parameters to filter out all transcripts without an open-reading frame and/or an open-reading frame shorter than 150 bp (Supplementary Table 10).

Genome guided assembly. We assembled a genome-guided assembly using the HiSat 2.0.5 – Stringtie v.1.3.2 suite[81]. We based our genome-guided assembly on the published guppy genome assembly (Guppy_female_1.0 + MT, RefSeq accession: GCA_000633615.1, latest release June 2016)[78]. Samples were individually mapped to the genome and built into transcripts using default parameters, but preventing the software from assembling de novo transcripts. The resulting individual assemblies were then merged into a single, non-redundant assembly using the built-in StringTie-merge function. In a similar fashion to the de novo assembly, we filtered out noncoding RNA and chose the best isoform for each transcript based on expression (Supplementary Table 10).

Reference transcriptome assembly. We used CD-Hit-Est to obtain a non-redundant reference transcriptome (RefTrans) by fusing the de novo and genome guided assemblies. Transcripts longer than 150 bp were clustered if they were >95% similar, preserving the longest representative for each cluster.

The resulting reference transcriptome was annotated by performing a BlastX to the non-redundant (nr) database of the NCBI. The associated gene identities obtained here were used to search multiple databases in all downstream GO annotations and pathway analysis as detailed below. See Supplementary Table 10 for details on the final number of transcripts preserved in the reference transcriptome and annotation statistics.

Differential expression. We quantified expression by mapping paired reads for each sample separately to the Reference Transcriptome using RSEM v.1.2.20[80], filtering transcripts <2 RPKM (reads per kilobase per million mapped reads), preserving only those transcripts that have expression values above this threshold in a least half of the samples for each treatment within a line. After this final filter, a total of 21,131 transcripts were kept for further analysis; 20,396 in the optic tectum and 19,571 in the telencephalon. Using sample correlations in combination with Multi-Dimensional Scaling (MDS) plots based on all expressed transcripts, we determined that out of the 54 samples, one optic tectum wild-type attractive male treatment sample, one optic tectum wild-type female treatment and one telencephalon small-brained female treatment sample were significant outliers and were therefore excluded from further analysis.

We relied on a random permutation test as described previously[4]. Filtered read counts were normalized using the standard function as implemented in DESeq2[82] (Supplementary Fig. 1) and used to perform a generalized linear model (GLM) to each transcript to evaluate the effect of treatment on expression level. Because we were interested in contrasting differences in expression associated with preference, we performed this analysis by grouping lines by their preference phenotype and carried out the GLM separately for Preference lines (wild-type and large-brained lines) and Non-preference lines (small-brained lines). After grouping samples by the female preference phenotype, the analysis was performed with six samples for the Preference lines and three samples for the Non-preference lines, except for treatments for which we had to remove one outlier (see Supplementary Table 11 for details on the sample sizes). In this way, we performed GLM to assess the significance of expression differences in pairwise comparisons between attractive and dull treatments, attractive and female treatments and, finally, dull and female treatments in Preference and Non-preference lines (Fig. 1b). To control for false positives and to determine which transcripts were differentially expressed between treatments, we used a random permutations test[4]. We generated 250 permuted datasets by randomly reassigning the sample names for the entire dataset of each tissue. Then we performed GLM in the exact same way as for the actual data, thus generating an empirical null distribution of 250 *P* values for each transcript. A transcript was considered differentially expressed when the statistic for the actual expression data fell below the 5% tail of the permutated data *P* value distribution. This method has been shown to better capture the structure of the data and does not assume independence across genes as other multiple test correction methods that can be over-corrective[4,83].

Our study relies on the assumption that messenger RNA levels correlate well with protein levels, which has been well supported in multiple other species[84–87]. Here, we used a differential expression approach so that the mRNA-to-protein ratio would be the same in all samples and therefore would not affect our results.

Differentially expressed genes involved in the mating decision: comparisons within Preference lines. To determine which genes are involved in the mating decision, we focused on the genes we found to be differentially expressed between the attractive and dull treatments in Preference lines. We applied several filters to the initial set of differentially expressed genes that passed the permutation threshold, retaining only those that have a potential role in mate choice based on their expression. We initially filtered out all genes that lack concordant expression (that is, genes that change in the same direction between pairs of treatments across all replicate samples) between attractive and dull treatments in all Preference lines, and then we retained those genes that are also differentially expressed between attractive and female treatments (Supplementary Fig. 1). Finally, we excluded genes that were also differentially expressed in dull male versus female comparisons, keeping only those genes associated with the evaluation of an attractive male (area P of Fig. 1b). Here, we assumed that any gene important in the evaluation of males of different qualities should also be differentially expressed between the attractive and female treatments. In this way, we were able to control for genes that change relative to social interaction alone. We refer to this final set of genes as Preference DE genes (Table 1).

Differentially expressed genes involved social interactions. We initially identified genes involved across the different social interactions we tested, independent of the female preference phenotype and the social context. For this purpose, we considered all genes determined to be differentially expressed across all three pairwise treatment comparisons separately within Preferences lines and Non-preference lines. These are genes that are differentially expressed in both mating contexts and general social interactions. Among these genes we selected only those that are differentially expressed in both Preference and Non-preference females, as these are the ones that become differentially expressed in different social contexts in all the guppies we studied, independent of their selection regime. We refer to these genes as Social DE genes.

Comparative analysis of genes involved in mate evaluation. To address the question of what genes and pathways differ between Preference and Non-preference females, we identified genes that were differentially expressed between attractive and dull treatments in Non-preference lines. We proceeded in the same fashion as described above for Preference DE genes (Non-preference DE genes; Table 1).

Co-expression networks. In order to study the relationship between genes expressed in the optic tectum and telencephalon, we used weighted correlation network analysis as implemented in the WGCNA package in R[43,44].

We built a weighted co-expression network for each tissue using genes that passed the expression filter described above and those with non-cero variance. In this way, we avoided using genes with non-significant variance and lowly expressed genes that generally represent transcriptional noise[43,44]. The input count data used to build co-expression networks was normalized and transformed using the variance-stabilizing transformation as implemented in DESeq2, as recommended by the authors of WGCNA. First, a Similarity matrix of the pairwise correlations between genes was built using log-transformed normalized data using a weighted combination of the Pearson correlation and Euclidean distance $S = \mathrm{SIGN}(\mathrm{corrx}) \times \{|\mathrm{corrx}| + [1 - \log(\mathrm{distx} + 1)]/\max[\log(\mathrm{distx} + 1)]/2\}$, as previously described[88]. We determined the most appropriate soft-threshold to use in order to reduce the number of spurious correlations based on the criterion of approximate scale-free topology[44], determined to be six for the telencephalon and four for the optic tectum. We used these soft-thresholds to build the Adjacency matrix and corresponding Topological Overlap matrix, a matrix of pairwise distance values between genes. Finally, we retained correlations >0.4, based on the correlation value distribution for each tissue, and genes that had >2 connections to other genes in the co-expression networks for all downstream analyses (Supplementary Fig. 4). Optic tectum and telencephalon network properties are summarized in Supplementary Table 3.

Module identification. We built a dendrogram of all genes based on the Topological Overlap matrix using hierarchical clustering in order to identify the gene modules in each tissue network. We then used the Dynamic Tree Cut method as implemented in WGCNA, using the 'tree' method and with a minimum cluster size of 30 genes, to detect the modules based on the clustering (Supplementary Fig. 5). The Dynamic Tree Cut method identified modules whose expression profiles are very similar. We did a further step to merge those modules with highly correlated expression values by estimating module eigengenes as described previously[43,44] (Supplementary Fig. 5).

Co-expression network analysis. Final co-expression networks were exported to Cytoscape[89] for further network data integration and visualization (Supplementary Fig. 1). Information on whether a gene was a differentially expressed gene or known to be a gene involved in social interaction and mate preference was attached to the network as metadata so that they could be visualized in all downstream network analyses (Fig. 4; Supplementary Fig. 4).

The Network Analyzer tool in Cytoscape was used to calculate network node attributes. These give an indication of how connected and central a gene is in the network. Here, we focused on three such attributes[90]: (1) Degree: the number of edges, that is other genes, each gene is connected to within the network. Central genes in the network will therefore have high degree values as opposed to more peripheral network genes. (2) Neighbourhood connectivity: defined as the average connectivity, or number of neighbours, for all its neighbours. (3) Clustering coefficient: the ratio of the number of edges between the neighbours of a gene, and the maximum number of edges that could possibly exist between such neighbours (number between 0 and 1). This is a measure of how connected a gene is relative to how connected it could be given the number of neighbours it has. This value will approach 0 for an unconnected gene and 1 for a fully connected gene in the centre of a network. We evaluated the connectivity and centrality of differentially expressed genes by examining the degree, neighbourhood connectivity and clustering coefficient of these genes in the optic tectum and telencephalon networks (Table 2; Supplementary Fig. 4). We carried out *t*-tests of log-transformed data to determine whether these attributes differ between the networks for the optic tectum and the telencephalon for each differentially expressed gene group (attractive versus dull in preference and Non-preference lines) and for gene groups known to be important in mating behaviour (lists in Supplementary Tables 4 and 5).

We performed enrichment tests to determine whether modules were enriched in differentially expressed genes of any category using one-tailed Fisher's exact tests (Fig. 3). We carried out similar tests to determine which modules in the network are enriched in genes previously known to be involved in social interactions and/or mate preference and in SPGs and IEGs.

Functional analyses. To study the biological functions and pathways associated with differentially expressed genes and gene modules, we obtained GO annotations for all expressed genes in the reference transcriptome that had a blast hit to the non-redundant and Swiss-Prot databases. We performed GO term enrichment tests in TopGO (R package) using the annotated reference transcriptome we built

as the background in one-tailed Fisher's exact tests with a threshold P value of <0.05 (Supplementary Table 6).

We determined which known pathways are associated with Preference DE genes within each module using hits to the human database in g:Profiler[62]. In a similar fashion, we investigated which TFs are known to regulate Preference DE genes within each module. This analysis was also based on data for humans, relying on the "TransFac transcription factor binding sites" database integrated into g:Profiler, as it is far more complete than databases for other species. Although providing a more complete view of the TF motifs associated with Preference DE genes, it is important to keep in mind that some TF motifs are likely to be different in a distant vertebrate like the guppy. Within TF motifs found to be enriched among Preference DE genes, we identified those for TFs with known roles in mate preference (Supplementary Table 4) as well as SPGs and IEGs (Supplementary Table 5). Additionally, we focused on TFs belonging to families previously identified in behavioural genetics studies, such as zinc finger proteins or POU domain TFs (Supplementary Fig. 8).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Normalized counts for all groups of differentially expressed genes as well as all expressed genes are available in Supplementary Data 1 and 2. RNA reads have been deposited at the NCBI Sequencing Read Archive, BioProject ID PRJNA413692. Additional data may be requested from the authors.

Received: 12 February 2018; Accepted: 29 August 2018;
Published online: 08 October 2018

References

1. Zayed, A. & Robinson, G. E. Understanding the relationship between brain gene expression and social behavior: lessons from the honey bee. *Annu. Rev. Genet.* **46**, 591–615 (2012).
2. 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).
3. Cummings, M. E. The mate choice mind: studying mate preference, aversion and social cognition in the female poeciliid brain. *Anim. Behav.* **103**, 249–258 (2015).
4. Ghalambor, C. K. et al. Non-adaptive plasticity potentiates rapid adaptive evolution of gene expression in nature. *Nature* **525**, 372–375 (2015).
5. Hitzemann, R. et al. Genes, behavior and next-generation RNA sequencing. *Genes Brain Behav.* **12**, 1–12 (2013).
6. Rosenthal, G. G. *Mate Choice. The Evolution of Sexual Decision Making from Microbes to Humans* (Princeton Univ. Press, Princeton, 2017).
7. Zahavi, A. Mate selection—a selection for a handicap. *J. Theor. Biol.* **53**, 205–214 (1975).
8. Kokko, H., Brooks, R., Jennions, M. D. & Morley, J. The evolution of mate choice and mating biases. *Proc. R. Soc. Lond. B* **270**, 653–664 (2003).
9. Robinson, G. E., Fernald, R. D. & Clayton, D. F. Genes and social behavior. *Science* **322**, 896–900 (2008).
10. Whitney, O. et al. Core and region-enriched networks of behaviorally regulated genes and the singing genome. *Science* **346**, 1256780 (2014).
11. Clayton, D. F. The genomic action potential. *Neurobiol. Learn. Mem.* **74**, 185–216 (2000).
12. Wang, S. M. T., Ramsey, M. E. & Cummings, M. E. Plasticity of the mate choice mind: courtship evokes choice-like brain responses in females from a coercive mating system. *Genes Brain Behav.* **13**, 365–375 (2014).
13. Cardoso, S. D., Teles, M. C. & Oliveira, R. F. Neurogenomic mechanisms of social plasticity. *J. Exp. Biol.* **218**, 140–149 (2015).
14. Cummings, M. E. et al. Sexual and social stimuli elicit rapid and contrasting genomic responses. *Proc. R. Soc. B* **275**, 393–402 (2008)..
15. Lynch, K. S., Ramsey, M. E. & Cummings, M. E. The mate choice brain: comparing gene profiles between female choice and male coercive poeciliids. *Genes Brain Behav.* **11**, 222–229 (2012).
16. Ramsey, M. E., Maginnis, T. L., Wong, R. Y., Brock, C. & Cummings, M. E. Identifying context-specific gene profiles of social, reproductive, and mate preference behavior in a fish species with female mate choice. *Front. Neurosci.* **6**, 62 (2012).
17. Wong, R. Y., Oxendine, S. E. & Godwin, J. Behavioral and neurogenomic transcriptome changes in wild-derived zebrafish with fluoxetine treatment. *BMC Genomics* **14**, 1 (2013).
18. Teles, M. C., Cardoso, S. D. & Oliveira, R. F. Social plasticity relies on different neuroplasticity mechanisms across the brain social decision-making network in zebrafish. *Front. Behav. Neurosci.* https://doi.org/10.3389/fnbeh.2016.00016 (2016).
19. Taborsky, B. & Oliveira, R. F. Social competence: an evolutionary approach. *Trends. Ecol. Evol.* **27**, 679–688 (2012).
20. Weitekamp, C. A. & Hofmann, H. A. Evolutionary themes in the neurobiology of social cognition. *Curr. Opin. Neurobiol.* **28**, 22–27 (2014).
21. Dukas, R. Evolutionary biology of animal cognition. *Annu. Rev. Ecol. Evol. Syst.* **35**, 347–374 (2004).
22. Woolley, S. C. & Doupe, A. J. Social context–induced song variation affects female behavior and gene expression. *PLoS Biol.* **6**, e62 (2008).
23. Houde, A. E. *Sex, Color, and Mate Choice in Guppies* (Princeton Univ. Press, Princeton, 1997).
24. Endler, J. A. Multiple-trait coevolution and environmental gradients in guppies. *Trends Ecol. Evol.* **10**, 22–29 (1995).
25. Brooks, R. Variation in female mate choice within guppy populations: population divergence, multiple ornaments and the maintenance of polymorphism. *Genetica* **116**, 343–358 (2002).
26. Houde, A. E. & Endler, J. A. Correlated evolution of female mating preferences and male color patterns in the guppy *Poecilia reticulata*. *Science* **248**, 1405–1408 (1990).
27. Endler, J. A. & Houde, A. E. Geographic variation in female preferences for male traits in *Poecilia reticulata*. *Evolution* **49**, 456–468 (1995).
28. Brooks, R. & Endler, J. A. Female guppies agree to differ: phenotypic and genetic variation in mate-choice behavior and the consequences for sexual selection. *Evolution* **55**, 1644–1655 (2001).
29. Sandkam, B., Young, C. M. & Breden, F. Beauty in the eyes of the beholders: colour vision is tuned to mate preference in the Trinidadian guppy *(Poecilia reticulata)*. *Mol. Ecol.* **24**, 596–609 (2015).
30. Hughes, K. A., Houde, A. E., Price, A. C. & Rodd, F. H. Mating advantage for rare males in wild guppy populations. *Nature* **503**, 108–110 (2013).
31. Rodd, F. H., Hughes, K. A., Grether, G. F. & Baril, C. T. A possible non-sexual origin of mate preference: are male guppies mimicking fruit? *Proc. R. Soc. Lond. B* **269**, 475–481 (2002).
32. Corral Lopez, A. et al. Female brain size affects the assessment of male attractiveness during mate choice. *Sci. Adv.* **3**, e1601990 (2017).
33. Kotrschal, A. et al. Artificial selection on relative brain size in the guppy reveals costs and benefits of evolving a larger brain. *Curr. Biol.* **23**, 168–171 (2013).
34. Chen, Y.-C. et al. Expression change in Angiopoietin-1 underlies change in relative brain size in fish. *Proc. R. Soc. B* **282**, 20150872 (2015).
35. Replogle, K. et al. The Songbird Neurogenomics (SoNG) Initiative: community-based tools and strategies for study of brain gene function and evolution. *BMC Genomics* **9**, 131 (2008).
36. Northcutt, R. G. Forebrain evolution in bony fishes. *Brain Res. Bull.* **75**, 191–205 (2008).
37. Bshary, R., Gingins, S. & Vail, A. L. Social cognition in fishes. *Trends Cogn. Sci.* **18**, 465–471 (2014).
38. Salas, C. et al. Neuropsychology of learning and memory in teleost fish. *Zebrafish* **3**, 157–171 (2006).
39. Derycke, S. et al. Neurogenomic profiling reveals distinct gene expression profiles between brain parts that are consistent in *Ophthalmotilapia* cichlids. *Front. Neurosci.* **12**, e1002962 (2018).
40. Lindholm, A. & Breden, F. Sex chromosomes and sexual selection in poeciliid fishes. *Am. Nat.* **160**, S214–S24 (2010).
41. Kirkpatrick, M. & Hall, D. W. Sexual selection and sex linkage. *Evolution* **58**, 683–691 (2004).
42. Kirkpatrick, M. & Ryan, M. J. The evolution of mating preferences and the paradox of the lek. *Nature* **350**, 33–38 (1991).
43. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* **9**, 559 (2008).
44. Zhang, B. & Horvath, S. A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* **4**, 17 (2005).
45. Iancu, O. D., Colville, A., Darakjian, P. & Hitzemann, R. Coexpression and cosplicing network approaches for the study of mammalian brain transcriptomes. *Int. Rev. Neurobiol.* **116**, 73–93 (2014).
46. Cummings, M. E. Sexual conflict and sexually dimorphic cognition—reviewing their relationship in poeciliid fishes. *Behav. Ecol. Sociobiol.* **72**, 73 (2018).
47. Galizia, G. & Lledo, P.-M. *Neurosciences—From Molecule to Behavior: A University Textbook* (Springer Science & Business Media, Berlin, 2013).
48. Langfelder, P. & Horvath, S. Fast R functions for robust correlations and hierarchical clustering. *J. Stat. Softw.* **46**, i11 (2012).
49. Jeong, H., Mason, S. P., Barabási, A. L. & Oltvai, Z. N. Lethality and centrality in protein networks. *Nature* **411**, 41–42 (2001).
50. Ramsey, M. E., Vu, W. & Cummings, M. E. Testing synaptic plasticity in dynamic mate choice decisions: *N*-methyl D-aspartate receptor blockade disrupts female preference. *Proc. R. Soc. B* **281**, 20140047 (2014).
51. Krumm, N., O'Roak, B. J., Shendure, J. & Eichler, E. E. A de novo convergence of autism genetics and molecular neuroscience. *Trends Neurosci.* **37**, 95–105 (2014).
52. Greco, B. et al. Autism-related behavioral abnormalities in synapsin knockout mice. *Behav. Brain Res.* **251**, 65–74 (2013).
53. Larhammar, D., Nordström, K. & Larsson, T. A. Evolution of vertebrate rod and cone phototransduction genes. *Phil. Trans. R. Soc. B* **364**, 2867–2880 (2009).

54. Moriguchi, S. et al. Reduced CaM kinase II and CaM kinase IV activities underlie cognitive deficits in NCKX2 heterozygous mice. *Mol. Neurobiol.* **21**, 1–12 (2017).
55. Cummings, M. E. & Ramsey, M. E. Mate choice as social cognition: predicting female behavioral and neural plasticity as a function of alternative male reproductive tactics. *Curr. Opin. Behav. Sci.* **6**, 125–131 (2015).
56. Wolf, C. & Linden, D. E. J. Biological pathways to adaptability—interactions between genome, epigenome, nervous system and environment for adaptive behavior. *Genes Brain Behav.* **11**, 3–28 (2012).
57. Cui, R., Delclos, P. J., Schumer, M. & Rosenthal, G. G. Early social learning triggers neurogenomic expression changes in a swordtail fish. *Proc. R. Soc. B* **284**, 20170701 (2017).
58. Okuyama, T. et al. A neural mechanism underlying mating preferences for familiar individuals in Medaka fish. *Science* **343**, 91–94 (2014).
59. Minatohara, K., Akiyoshi, M. & Okuno, H. Role of immediate-early genes in synaptic plasticity and neuronal ensembles underlying the memory trace. *Front. Mol. Neurosci.* **8**, 78 (2015).
60. Cummings, M. E. Looking for sexual selection in the female brain. *Phil. Trans. R. Soc. B* **367**, 2348–2356 (2012).
61. Kowiański, P. et al. BDNF: a key factor with multipotent impact on brain signaling and synaptic plasticity. *Cell. Mol. Neurobiol.* **38**, 579–593 (2018).
62. Reimand, J. et al. g:Profiler-a web server for functional interpretation of gene lists (2016 update). *Nucleic Acids Res.* **44**, W83–W89 (2016).
63. Herbert, J. Peptides in the limbic system: neurochemical codes for co-ordinated adaptive responses to behavioural and physiological demand. *Prog. Neurobiol.* **41**, 723–791 (1993).
64. O'Connell, L. A. & Hofmann, H. A. The vertebrate mesolimbic reward system and social behavior network: a comparative synthesis. *J. Comp. Neurol.* **519**, 3599–3639 (2011).
65. O'Connell, L. A. & Hofmann, H. A. Evolution of a vertebrate social decision-making network. *Science* **336**, 1154–1157 (2012).
66. Alexander, H. J., Taylor, J. S., Wu, S. S. T. & Breden, F. Parallel evolution and vicariance in the guppy (*Poecilia reticulata*) over multiple spatial and temporal scales. *Evolution* **60**, 2352–2369 (2006).
67. Suk, H. Y. & Neff, B. D. Microsatellite genetic differentiation among populations of the Trinidadian guppy. *Heredity* **102**, 425–434 (2009).
68. Kotrschal, A., Corral Lopez, A., Amcoff, M. & Kolm, N. A larger brain confers a benefit in a spatial mate search learning task in male guppies. *Behav. Ecol.* **26**, 527–532 (2015).
69. van der Bijl, W., Thyselius, M., Kotrschal, A. & Kolm, N. Brain size affects the behavioural response to predators in female guppies (*Poecilia reticulata*). *Proc. R. Soc. B* **282**, 20151132 (2015).
70. Kotrschal, A., Kolm, N. & Penn, D. J. Selection for brain size impairs innate, but not adaptive immune responses. *Proc. R. Soc. B* **283**, 20152857 (2016).
71. Kotrschal, A., Corral Lopez, A., Szidat, S. & Kolm, N. The effect of brain size evolution on feeding propensity, digestive efficiency, and juvenile growth. *Evolution* **69**, 3013–3020 (2015).
72. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).
73. Corral Lopez, A., Eckerström-Liedholm, S., Der Bijl, W. V., Kotrschal, A. & Kolm, N. No association between brain size and male sexual behavior in the guppy. *Curr. Zool.* **61**, 265–273 (2015).
74. Corral Lopez, A., Garate-Olaizola, M., Buechel, S. D., Kolm, N. & Kotrschal, A. On the role of body size, brain size, and eye size in visual acuity. *Behav. Ecol. Sociobiol.* **71**, 179 (2017).
75. Kotrschal, A. et al. Brain size does not impact shoaling dynamics in unfamiliar groups of guppies (*Poecilia reticulata*). *Behav. Processes* **147**, 13–20 (2018).
76. Kotrschal, A. et al. Evolution of brain region volumes during artificial selection for relative brain size. *Evolution* **71**, 2942–2951 (2017).
77. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
78. Künstner, A. et al. The genome of the Trinidadian guppy, *Poecilia reticulata*, and variation in the Guanapo population. *PLoS ONE* **11**, e0169087 (2016).
79. Grabherr, M . G. et al. Trinity: reconstructing a full-length transcriptome without a genome from RNA-seq data. *Nat. Biotechnol.* **29**, 644–652 (2011).
80. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* **12**, 323 (2011).
81. 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. *Nat. Protoc.* **11**, 1650–1667 (2016).
82. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
83. Slonim, D. K. From patterns to pathways: gene expression data analysis comes of age. *Nat. Genet.* **32 Suppl.**, 502–508 (2002).
84. Dean, R. & Mank, J. E. Tissue specificity and sex-specific regulatory variation permit the evolution of sex-biased gene expression. *Am. Nat.* **188**, E74–E84 (2016).
85. Greenbaum, D., Colangelo, C., Williams, K. & Gerstein, M. Comparing protein abundance and mRNA expression levels on a genomic scale. *Genome Biol.* **4**, 117 (2003).
86. Futcher, B., Latter, G. I., Monardo, P., McLaughlin, C. S. & Garrels, J. I. A sampling of the yeast proteome. *Mol. Cell. Biol.* **19**, 7357–7368 (1999).
87. Lundberg, E. et al. Defining the transcriptome and proteome in three functionally different human cell lines. *Mol. Syst. Biol.* **6**, 450 (2010).
88. Inbar, E. et al. The transcriptome of Leishmania major developmental stages in their natural sand fly vector. *mBio* **8**, e00029–17 (2017).
89. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
90. Maslov, S. & Sneppen, K. Specificity and stability in topology of protein networks. *Science* **296**, 910–913 (2002).
91. Suzuki, R. & Shimodaira, H. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* **22**, 1540–1542 (2006).

Acknowledgements

This work was funded by a Marie Sklodowska-Curie Fellowship (654699) and a National Science Foundation Postdoctoral Fellowship in Biology (1523669) to N.I.B., by grant agreements 260233 and 680951 from the European Research Council to J.E.M., a Swedish Research Council grant (2016-03435) to N.K. and a Knut and Alice Wallenberg grant (102 2013.0072) to N.K. We gratefully acknowledge support from a Royal Society Wolfson Merit Award to J.E.M. We thank P. Almeida, I. Darolti, J. Morris, V. Oostra, A. Wright and T. Price for valuable discussions and help with manuscript preparation. We thank the Oxford Genomics Centre at the Wellcome Centre for Human Genetics (funded by a Wellcome Trust grant (reference 203141/Z/16/Z)) for the generation and initial processing of the sequencing data, and the UCL Legion High Performance Computing Facility (Legion@UCL).

Author contributions

N.I.B., A.C.-L., N.K. and J.E.M. conceived of the study and designed the experiments. A.K. and N.K. created the brain size selection lines. A.K. and S.D.B. performed laboratory work for fish housekeeping. A.C.-L. and S.D.B. selected fish for the experiments. A.C.-L. performed the behavioural tests and dissected the brain regions. N.I.B. performed all laboratory RNA work and analysed the data. All authors contributed to writing the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41559-018-0682-4.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to N.I.B.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature research

Corresponding author(s): Natasha I. Bloch

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	NA
Data analysis	Quality control: FastQC. v.0.11.4. 553 (www.bioinformatics.babraham.ac.uk/projects/fastqc) de novo assembly: Trinity v2.2, RSEM v1.2.20, Transdecoder (Transdecoder v3.0.1, http://transdecoder.github.io). Genome-guided assembly: HiSat 2.0.5 - Stringtie v1.3.2 suite Refrence transcriptome costruction: CD-Hit-Est Differential expression: RSEM v1.2.20, DESeq2

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Normalized counts for all groups of differentially expressed genes as well as all expressed genes are available as Supplementary Datasets. RNA reads have been deposited at the NCBI Sequencing Read Archive, BioProject ID PRJNA413692. Additional data may be requested from the authors.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	270 samples total, divided into 54 pools (27 pools for each the optic tectum and the telencephalon). We performed our experiment in 3 lines (WT, LB and SB), conducting 3 treatments for each line. For each line/treatment combination in our experiment, we had 15 replicate females, divided into 3 pools. This is equivalent to 9 pools for each line and 9x3=27 pools, with 5 female brains per pool, 135 females. For each female we dissected the optic tectum and telencephalon separately for sequencing.
Data exclusions	Some samples were clear outliers when considering overall expression and were thus excluded from the analysis. These samples included 2 optic tectum samples and 1 telencephalon sample as detailed in the supplementary materials.
Replication	For each treatment within each line, we have 3 pools composed of 5 individual females each (135 females total). Moreover we have used two separate lines with the same preference phenotype corresponding to 2 independent biological replicates for the preference phenotype.
Randomization	Within each line, focal females were chosen at random and randomly matched with the different presentation males. Focal females belonging to same replicate selection line and the same treatment were presented with different males to avoid uncontrolled male-driven changes in expression. For this, we balanced the number of large-brained, small-brained and wild-type females presented to colorful males, dull males and females respectively per day
Blinding	Investigators were blind to group allocation and data collection. Moreover, trials were videotaped so researchers were not in the room while behavioral trials were conducted to avoid interference.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Guppies used in our experiment are laboratory-raised descendants of Trinidad guppies sampled from the high predation populations of the Quare River (Trinidad). We based our study on guppies from this wild-type population and two different selection lines, derived from the wild-type fish, which have been selected on relative brain size. In summary fish were indirectly

selected based on parental brain size achieving a difference of up to 13.6% in relative brain size between lines selected to have small brains, here denoted small brain lines (SB lines), and lines selected to have larger brain (LB lines). Three replicate lines were generated for each selection regime.

FOCAL FEMALES: For this study, we used virgin females from the fifth generation of selection, all aged approximately 6 months. We used a total of 135 females, 45 from each line.
At the end of each trial females were euthanized by transfer to ice water. After 45 seconds, and with aid of a Leica S4E microscope, we removed the top of the skull to expose the brain.

PRESENTATION MALES and FEMALES: For our study we divided females among 3 treatments: two treatments represented a male evaluation context, in which females were presented either an attractive male (attractive treatment) or an unattractive male (dull treatment), and a third treatment in which females were exposed to another female representing a general social interaction treatment. We used 30 presentation males selected for their colorful or dull patterns. Presentation females were chosen at random from the wild-type population. Presentation animals were reintroduced into their original laboratory populations after teh experiments.

The experiment was performed in accordance with ethical applications approved by the Stockholm Ethical Board (Reference number: N173/13, 223/15 and N8/17). These applications are consistent with the Institutional Animal Care and Use Committee guidelines

Wild animals

The study did not involve wild animals

Field-collected samples

The study did not involve field-collected samples