Mate choice in the brain: Species differ in how male traits 'turn on' gene expression in female brains

Jason Keagy[1,2]*, Hans A. Hofmann[3], and Janette W. Boughman[2]

[1]Department of Ecosystem Science and Management, The Pennsylvania State Pennsylvania, University Park, PA, USA

[2]Department of Integrative Biology, Michigan State University, East Lansing, MI, USA

[3]Department of Integrative Biology, Institute for Neuroscience, The University of Texas at Austin, Austin, TX, USA

* Corresponding author: Jason Keagy; Email: keagy@psu.edu

Abstract

Mate choice plays a fundamental role in speciation, yet we know little about the molecular mechanisms that underpin this crucial decision-making process. Stickleback fish differentially adapted to limnetic and benthic habitats are reproductively isolated and females of each species use different male traits to evaluate prospective partners and reject heterospecific males. Here, we integrate behavioural data from a mate choice experiment with gene expression profiles from the brains of females actively deciding whether to mate. We find substantial gene expression variation between limnetic and benthic females, regardless of behavioural context, suggesting general divergence in constitutive gene expression patterns, corresponding to their genetic differentiation. Intriguingly, female gene co-expression modules covary with male display traits but in opposing directions for sympatric populations of the two species, suggesting male displays elicit a dynamic neurogenomic response that reflects known differences in female preferences. Furthermore, we confirm the role of numerous candidate genes previously implicated in female mate choice in other species, suggesting that evolutionary tinkering with these conserved molecular processes underlies divergent mate preferences and sexual isolation. Taken together, our study adds important new insights to our understanding of the molecular processes underlying female decision-making critical for generating sexual isolation and speciation.

Keywords: mate choice, gene expression, brain, speciation, stickleback

Introduction

Choosing a mate is a key fitness decision [1] often crucial to speciation [2–4]. Much research has sought to identify the displays individuals assess when making mate choice decisions [5,6] as well as understand the evolution and divergence of mate preferences [4,7,8]. However, we know little about the cognitive and molecular mechanisms that underpin such decision-making or how these mechanisms vary across species [9,10]. Comparative transcriptomic studies, where constitutive gene expression profiles are systematically analyzed across different populations, species, and environments, have already provided important insights into the evolution of molecular mechanisms underlying various complex behaviors, such as learned vocalizations [11], mating systems [12,13], and cooperation [14]. Recent studies have also identified dynamic transcriptomic variation related to mate choice [15,16]. Yet whether and how the neuromolecular mechanisms reflecting these decision-making processes change as mate preferences diverge in the process of speciation has not been examined, nor do we know how variation in gene expression contributes to reproductive isolation. Given the importance of sexual selection to both evolution and speciation, this is a critical gap. To answer these questions, we urgently need studies that examine gene expression differences in individual brains as they choose whether to mate with conspecifics and heterospecifics.

Neural transcriptomes exhibit both constitutive (e.g., dependent on species or sex) and dynamic (e.g., determined by behavioural state) components, which makes them uniquely suited to gain insight into the molecular processes underlying behavioural diversification and speciation [13,17,18]. We ask here how female brain gene expression varies among closely-

related species, focusing on transcriptomic responses during courtship. We study threespine stickleback fish (*Gasterosteus aculeatus*). The limnetic-benthic threespine stickleback species-pairs show parallel phenotypic divergence and speciation; each species has evolved independently in multiple lakes, showing repeated and substantially parallel divergence from ancestral marine fish [19,20] confirming that both sexual selection and natural selection contribute [20–24]. Limnetic and benthic species experience strong sexual isolation, but mate freely with their own species whether from their own or another lake [22]. Moreover, prior research has revealed that the species have diverged in female preferences for nuptial colour, odour, courtship behaviour, body shape and size; and females reject heterospecific males based on differences in these traits [23,25–29]. Finally, limnetic females are more responsive to male courtship and have stronger conspecific preferences than benthic females [26,30].

In our study, limnetic and benthic females were collected from two lakes where the limnetic-benthic divergence is well-studied and thought to be evolutionarily independent to provide two evolutionary replicates for testing parallel patterns of gene expression. Once in the lab, females were courted either by a conspecific or heterospecific male from their lake or were placed with a conspecific female from their lake as a social control (Fig. 1A). We quantified male morphological and behavioural display traits, female courtship behaviours, and female preference. Because the sample sizes were very small for one of the benthic populations, we generated whole brain transcriptomes from females from two limnetic and one benthic population (see also Materials and Methods). Finally, we evaluated the relationship of these neural transcriptomes with behavioural and morphological data from the behavioural

experiment. Although our transcriptome data ended up being unbalanced, we can still compare limnetic populations to each other to test for parallel changes among them and we can compare limnetic and benthic fish from the same lake to test for divergence in gene expression as part of the speciation process.

We make the following predictions for our transcriptome data. First, given the divergent evolution of benthic and limnetic species in both ecology and reproduction, we predict divergence between the species in constitutive gene expression patterns. Second, we predict limnetic fish from the two lakes will show similar patterns of constitutive gene expression because of their parallel evolution. Third, we predict females will show dynamic brain gene expression differences depending on the social context (conspecific courtship, heterospecific courtship, or social control) as they will not only be experiencing different stimuli, but also making different decisions with respect to reproduction. Given that benthic and limnetic females have divergent preferences and are influenced by different male traits, in particular nuptial colour, courtship behaviour, odour, and body size and shape [23,25–29], we predict that the genes implicated in conspecific preference and social context might vary, especially when comparing benthic and limnetic fish from the same lake where reinforcement is expected to have occurred. Alternatively, these dynamic attributes of the brain transcriptome may be similar between benthic and limnetic species due to being dependent on conserved molecular pathways. Likewise, we predict limnetic fish from the two different lakes will show similar patterns of dynamic gene expression as a consequence of parallel evolution. Alternatively, the

parallel evolution of behaviour may not be underpinned by parallel genomic or transcriptomic processes [31–33].

Finally, we predict that gene expression in female brains will respond to variation in male displays in a quantitative manner. What we envision here is that gene expression levels will increase (or decrease) in a linear manner when the display is more attractive, and the opposite will occur when the display is disliked. Because females of the two species either accept or reject males based on distinct display traits, their gene expression responses should be in opposing directions, in other words, up- or down-regulated in proportion to the trait value experienced. Another possibility is that different traits might activate similar gene expression patterns, with a given trait only activating expression in the relevant species that uses that trait for mate choice. Both possibilities would be reflected by a significant species-by-trait interaction in a linear model analysis. The absolute direction of response for each species is difficult to predict because changes in gene expression may ultimately lead to activation or inhibition of behaviour. Some of the strongest evidence linking female mating decisions to male displays and reproductive isolation in a neuromolecular manner would be to find differential gene expression patterns that are associated with *both* female choice behaviour and male display as this would imply they are connected mechanistically.

Materials and Methods

Benthic and limnetic threespine stickleback were collected from two lakes on Texada Island, British Columbia (BC, Fig. 1) and transferred to Michigan State University. Additional fish care information can be found in the Supplemental Material.

Behavioural trials

We used standard methods for assessing female mate preference in stickleback in which a female is courted by a single male [25,26,34]. Each female subject experienced one of three trial conditions: 1) courtship by a male from their lake and species (conspecific courtship, n = 29), 2) courtship by a male from their lake and the opposite species (heterospecific courtship, n = 28), or 3) a control in which female subjects were with another female from her same lake and species (social control, n =26) (Fig. 1A).

During May-August 2014, reproductive males were taken from holding aquaria and placed individually in new visually isolated aquaria with nesting materials and enticed to build a nest (see Supplementary Material for more details). All behavioural trials commenced between 0930 ET and 1230 ET. We verified each female's reproductive status by gently squeezing her abdomen to confirm presence of ripe eggs [27]. For the courtship treatments, females were placed in an opaque holding container just below the water surface in the male's nesting aquarium for a 5 min acclimation period. The female was then remotely released, and the behaviours of both male and female were recorded using the event recorder JWatcher 1.0 (http://www.jwatcher.ucla.edu/) by two observers. The social control treatment was done

similarly except that the stimulus fish was a non-reproductive female from the subject's home aquarium.

After 20 min or spawning (whichever came first, see Supplementary Material for more details), the female subject was killed by rapid cervical dissection, and the brain was immediately dissected under a microscope in a Sylgaard 184-lined petri-dish filled with a Ringer's solution. Brains were stored overnight in RNAlater at 4°C and then transferred to -20°C until RNA extraction.

RNA-seq

In September 2014, brains were shipped on dry ice to The University of Texas at Austin where RNA was extracted using the Maxwell 16 LEV simplyRNA Tissue Kit, which utilizes a robot to increase consistency. At the UT Austin Genomic Sequencing and Analysis Facility (GSAF) each sample (all with RIN > 7.8, 8.8 ± 0.4 mean ± SD, 11 samples do not have RIN scores) was prepared for RNA-seq using Poly-A mRNA capture and given a unique barcode. A single library of all multiplexed samples was sequenced across eight lanes of an Illumina Hiseq 2500 with 2x50 PE chemistry. Sample size for each treatment for Paxton limnetic, Paxton benthic, and Priest limnetic populations was n = 5-6. Brains were randomly chosen for each population and treatment if there were >6 to choose from. Very few Priest benthic females became reproductive, resulting in three or fewer brains per treatment for this population. Given that we were expecting large individual variation in transcriptomes, we focused our sequencing effort

on the other three populations with larger sample sizes. Bioinformatic analyses were carried out using the computational resources of the Texas Advanced Computing Center (TACC).

RNA-seq resulted in 34.2 ± 4.6 million reads (mean ± SD) per sample per sequencing direction (forward or reverse). Reads from all eight lanes were combined for each uniquely barcoded individual, separately for forward and reverse reads. We conducted quality control checks using *fastqc* (v0.11.1). We then ran *Trimmomatic* (v0.33) to remove a small amount of adapter contamination. Next, reads were aligned to the *Gasterosteus aculeatus* Ensembl BROAD S1 draft genome (version 78) using *bwa* (v0.7.7). We obtained expression information for 21,798 genes (Supplementary Table 1). *Samtools* (v1.2) was used to convert sam to bam files, and then sort and index them. These sorted and indexed bam files were then passed to *bedtools* (v2.23.0) to count gene transcripts. The resulting gene counts were analysed quantitatively using *DESeq2* [35] and *WGCNA* [36,37] in *R* v4.2.2 [38] as described below.

Analysis

We derived two principal components each for variation in female behaviour (Fbehav), male behaviour (Mbehav), and male morphology (Mmorph) to facilitate the integration of transcriptome data with female mating behaviour and the male traits females experienced (Supplementary Fig. 1). Derivation and analyses of female and male phenotypic data are described in the Supplementary Material.

Using the *rlog* function in the *DESeq2* library [35] and a design matrix with an intercept only, we first applied a "regularized log" transformation to the gene expression count matrix which was exported for the WGCNA and candidate gene analyses (Supplementary Table 10, more below). Using these normalized and variance-stabilized gene expression counts, we conducted a principal components analysis on the 90% most variable genes as an initial step to visualize how populations and treatments separate along the axes of greatest gene expression variation. Two samples appeared to have been swapped and were removed from analyses (Fig. 2A). The original untransformed gene expression count matrix was then analysed using standard methods [35] with a model where gene expression counts were predicted by the independent effects of treatment and population. To test for constitutive gene expression differences between populations, subsequent contrasts were computed using the "normal" shrinkage option and an alpha of 0.1. To determine genes involved in conspecific preference or social context, treatment difference contrasts were calculated separately for each population from a second model in which gene expression counts were predicted by the interactive effect of treatment and population. The *gprofiler2* library was used for GO analysis to identify the biological processes, cellular locations, and molecular functions associated with differentially expressed genes [39].

To identify gene modules that may underpin shared biological functions, we conducted weighted gene co-expression network analysis (WGCNA) using standard methods for signed network construction using the *WGCNA* library [36,37]. Briefly, we used the 90% most variable genes from the normalized and variance-stabilized gene expression count matrix (without the

two samples suspected of being swapped). Additional quality control steps eliminated two samples that had the poorest read-mapping percentage. Using the TACC high-performance computing cluster, we tried different combinations of parameters to optimize network construction rather than simply relying on defaults. The analysis presented here used the following parameters: maxBlockSize = 30,000, power = 9, network-type = "signed", corType = "bicor", maxPOutliers = 0.05, minModuleSize = 30, mergeCutHeight = 0.25, deepSplit = 2 (default), and detectCutHeight = 0.995 (default). Modules were analysed first using t-tests to compare sympatric species (Paxton limnetic vs. Paxton benthic), limnetic species from different lakes, and treatments. Then we constructed linear models (using the *lm* function in the *stats* library [38]) with module eigengene expression predicted by population, morphological or behavioural trait, and their interaction. We assessed statistical significance of main and interaction effects using an empirical FDR procedure [40]. Briefly, significance was initially calculated using an Analysis of Variance table with type-II sums of squares from the linear model results (using the *Anova* function in the *car* library [41]). Then the gene expression counts were shuffled among samples 10,000 times and new significance values calculated each time. eFDR was the proportion of times the true significance value was lower than or equal to the significance values calculated from these shuffled datasets. The *gprofiler2* library was used for GO analysis to identify the biological processes, cellular locations, and molecular functions associated with different modules [39].

We also compiled a list of candidate genes from the existing literature that had previously been implicated in mate choice and analysed their expression using linear models with gene

226 expression predicted by population, morphological or behavioural trait, and their interaction
227 and significance assessed using an empirical FDR procedure.

228

229 **Results and Discussion**

230 ***Females preferred conspecific males and males differed in species-specific ways***

231 Consistent with earlier work [22,26,42], females displayed stronger behavioural preference for
232 conspecific males than heterospecific males (ANOVA: $F_{1,49} = 11.11$, $P = 0.016$, Fig. 1B). Also
233 consistent with earlier work, limnetic females showed especially strong conspecific preference
234 (ANOVA: $F_{1,36} = 16.90$, $P = 0.0062$). Moreover, male morphological and behavioural display
235 traits differed between the species (Fig. 1C).

236

237 ***Species show highly differentiated gene expression for individual genes***

238 We first asked whether constitutive variation in female brain transcriptomes is concordant with
239 previously described genetic differences between benthic and limnetic fish [19,43–46]. We
240 found that gene expression patterns were indeed highly differentiated between the one
241 benthic population and two limnetic populations (Fig. 2; Supplementary Fig. 2). Unexpectedly,
242 benthic and limnetic fish from the same lake had fewer differentially expressed genes, DEGs,
243 than benthic and limnetic fish from different lakes (7,443 [34% total genes] vs 8,370 [38%],
244 adjusted p-value < 0.1, 21,796 total genes; Fig. 2B; Supplementary Table 2), possibly due to
245 lake-specific ecology or low levels of past or contemporary gene flow [30] partly homogenizing
246 genetic differences underlying constitutive expression patterns. The two limnetic populations

had similar overall expression patterns (more overlap in scores of the first principal component describing multivariate gene expression variation and fewer DEGs, Fig. 2), consistent with previously described parallel evolution and speciation [22,47,48]. Nonetheless, female brains of the Paxton and Priest limnetic fish did show differences in gene expression (3,642 [17%] DEGs, adjusted p-value < 0.1, 21,796 total genes; Fig. 2B; Supplementary Fig. 2C; Supplementary Table 2) and different gene expression patterns in response to specific male traits (see below).

Magnitude of gene expression differences between treatments reflect species differences in strength of courtship responsiveness and conspecific preference

We euthanized all fish within 20 min stimulus onset to reveal dynamic transcriptome activity during decision-making. Even with this short stimulus time, limnetic females from both lakes showed numerous DEGs in comparisons between the conspecific male treatment versus either the conspecific female (social control) or heterospecific male treatments (28-52 [0.13-0.24%], adjusted p-value < 0.1, 21,796 total genes; Fig. 3; Supplementary Table 3). Benthic females, in contrast, had fewer DEGs for both of these comparisons (7, Fig. 3), consistent with prior behavioural work indicating limnetic females are more responsive to male courtship and have stronger conspecific preferences than benthic females [26,48]. These patterns are interesting in light of earlier findings in the Panuco swordtail (*Xiphophorus nigrensis*), where females strongly prefer large courting males to small coercive males [49]. This robust behavioural preference in swordtails was reflected in the brain transcriptome: the brains of females in this choice situation have many more DEGs than females exposed to only small males or only females [15]. Recent work in another poecilid species, the guppy (*Poecilia reticulata*), also demonstrated that

females that expressed a strong preference for colourful over drab males had many more DEGs than females that did not have a male coloration preference [16]. Taken together with our findings, these results point to stronger gene expression response with greater mate preference.

Genes involved in conspecific preference are population-specific

DEGs from the comparison of conspecific male to heterospecific male treatments ("conspecific preference", 7-46 DEGs, [0.03-0.21%], adjusted p-value < 0.1, 21,796 total genes; Fig. 3C; Supplementary Table 3) support the conclusion that female brains are differentiating conspecific from heterospecific males and that dynamic changes in gene expression are involved in this decision-making. If female gene expression was not involved in discriminating between conspecific and heterospecific males, we would expect no DEGs in this comparison. Instead, our results suggest that conspecific and heterospecific males elicit differential expression of certain genes in female brains. Notably, however, we found virtually no overlap in DEGs between populations for conspecific preference (Fig. 3C), suggesting that different genes are involved in conspecific preference in each population. We expected this for benthic compared to limnetic populations because they have distinct mating preferences and social behaviour [26,48,50]; our findings support that prediction. However, we expected some overlap between limnetic populations because of their parallel evolution in ecological and reproductive traits [26,48,50] and because of their overall similarity in constitutive gene expression discussed above. Instead, the little overlap between limnetic populations for

conspecific preference is consistent with independent behavioural evolution in each lake [19] resulting in distinct molecular networks related to conspecific preference (reflected also by non-overlap of Gene Ontology (GO) terms, Supplementary Table 4). This could have been either due to the available genetic variants differing (*sensu* the mutation order hypothesis [51–53]), or because selection was less similar than thought. Although there are exciting examples of parallel gene expression changes underlying parallel evolution of complex behavioural phenotypes [11–14,33], it seems that parallel phenotypic evolution often rests on only partly parallel genetic mechanisms [31,54], as is suggested by the data from the two limnetic populations here.

Gene co-expression network analysis supports individual gene analysis results

We next used weighted gene co-expression network analysis (WGCNA [36,37]) to identify gene modules that may underpin shared biological functions (Supplementary Fig. 3, Supplementary Table 5); expression of genes in a module can be summarized by eigengenes, the first principal component of a given module. We identified 13 differentially expressed module eigengenes (DEMEGs) between limnetic and benthic female brains from the same lake (Fig. 4; Supplementary Table 6), similar to patterns we uncovered with the individual gene analyses. Also similarly, we found fewer DEMEGs when comparing limnetic females from different lakes (7 DEMEGs; Fig. 4; Supplementary Table 6). However, five of these seven modules also differentiate Paxton limnetic from Paxton benthic fish and so separate populations generally. Thus, we find support for our prediction of stronger divergence in constitutive gene expression

between diverged species and weaker divergence between populations of limnetic fish that have evolved in parallel.

Although we found relatively few DEMEGs between treatments (Fig. 4; Supplementary Table 6), we did identify two DEMEGs when females were courted by a conspecific versus heterospecific male; one for Paxton benthic and one for Paxton limnetic females. We suggest that these modules are involved in conspecific preference and may play an important role in sexual isolation. We also found three DEMEGs when benthic females were courted by a conspecific male versus interacted with a female, suggesting these modules are involved in evaluating males. We were surprised that benthic females showed a strong effect here as they neither exhibited strong preference nor strongly discriminated between males at the behavioural level (Fig. 1B) and they had the fewest DEGs for these comparisons (Fig. 3; Supplementary Table 3). Module eigengene expression likely reflects female perception and decision-making rather than a final decision given the relatively short time (≤20 min) between trial start and sampling brains. We therefore examined next how female module eigengene expression relates to male display trait variation.

Module eigengene expression in the female brain reflects individual variation in a male sexually selected trait in population-specific ways

It is well established that variation in male traits influences both current and future female mating decisions [1,5,6,34], the specific traits that females focus on vary between populations

[55,56], and changes in the expression and topology of gene networks interface with neural networks to influence future behaviour [57]. We therefore predicted that module eigengene expression in the female brain should vary dynamically in conjunction with male displays, likely in a species- or population-dependent fashion. Indeed, this was true for an important male morphological trait, throat colour, which varies between benthic and limnetic species, is involved in female choice and male competition differently for the two species, and is critical for sexual isolation [23,24,58–60]. Variation in male throat colour significantly predicted variation in the eigengene expression of two modules in a population-dependent manner (Fig. 5; Supplementary Table 7), with mean expression and slopes differing between the benthic population and at least one limnetic population, especially the sympatric one. Thus, the activity of brain gene co-expression modules was dynamically altered in female brains in response to this key male trait known to be subject to sexual selection and involved in sexual isolation. Importantly, the relationship between module eigengene expression and male throat colour was strongly divergent in benthic and limnetic females from the same lake, in parallel with strongly divergent preferences for this trait in these same females [23]. This finding argues that the differential recruitment of gene networks implied by these co-expression modules underpins divergence in mating behaviour and conspecific preference, thus contributing to reproductive isolation between diverging species.

Our finding that module eigengene expression responds to male throat colour helps to explain the low number of DEMEGs for treatment comparisons. Based on these results, females in the same treatment would be expected to have variable expression of some modules based on the

trait values of the specific males courting them. Furthermore, although limnetic males tend to have more throat coloration than benthic males, there is overlap in the distribution, leading to imperfect correspondence between coloration and treatment, likely obscuring the gene expression signal from treatment comparisons.

Moving from individual traits to multivariate descriptions of male and female phenotypes

Mate choice decisions are typically based on more than a single trait [10,55,61,62]. We therefore used the axes of variation in male behaviour (Mbehav PC1 and PC2) and morphology (Mmorph PC1 and PC2) as well as female behaviour (Fbehav PC1 and PC2) we previously inferred using PCA (Supplementary Fig. 1) to integrate variation in these traits with our gene expression data. We did this using linear models, testing for population and trait PC main effects as well as their interactive effects on module eigengene expression (Fig. 6; Supplementary Table 7). Significant main effects of trait PCs (with no population-by-trait interaction) indicate modules that respond to specific trait PCs in a consistent manner across all populations. Genes in these modules could reflect conserved patterns of gene expression during courtship. We did find this pattern with the PC describing male size and body coloration (Mmorph PC2) for 4 modules, the PC describing male courtship vigour (Mbehav PC1) for 3 modules, and the PC describing the nature of female response (Fbehav PC2) for 1 module, for a total of 5 unique modules. Several of these modules have clearly differentiated biological functions, according to GO analysis (Supplementary Table 8), that suggest involvement of stress, metabolism, neural function, and DNA/RNA processing molecular pathways.

Most importantly, however, significant population-by-trait interactions address our primary prediction that divergent selection on female mate preferences would result in divergent brain gene expression patterns in response to male display trait values, especially in benthic and limnetic females from the same lake. Indeed, we saw many modules with this pattern; 7 modules showed a significant population-by-trait interaction and, in all cases, limnetic and benthic females from the same lake had opposing slopes (Fig. 6). When we shifted our focus to female behaviour in response to courtship, we found a similarly large number of modules (n = 8) that showed a population-by-trait interaction, specifically the nature of female response (early vs. late, Fbehav-PC2). Once again, we observed that limnetic and benthic females from the same lake had opposite slopes, showing strong divergence in gene expression in association with female behaviour.

GO analysis of the 8 modules with a significant population-by-Fbehav-PC2 interaction indicates distinct biological functions being overrepresented by module genes (Supplementary Fig. 4; Supplementary Table 8), especially for the greenyellow and green modules: DNA/RNA processing, metabolic processes, and cellular stress response (greenyellow) and synaptic signalling, ion transport, cell-cell signalling, and neural development (green). Clearly something very different is happening in female brains of sympatric species of stickleback. The preponderance of genes with neural functions (green module) and those influencing future gene expression and response to stress (greenyellow module) suggests those modules are mediating activity in the brain involved in decision-making. The different expression patterns

for limnetic and benthic females from the same lake point to a key role for these modules in isolating the species.

Even more interesting, for five modules there was overlap where a module's eigengene expression was associated with *both* male trait and female behavioural variation (Fig. 6). We interpret this to indicate that the male trait elicits a specific neurogenomic response in the female brain that, in turn, influences female choice behaviour. For example, the red module showed lower eigengene expression in Paxton limnetic females when they were courted by males that they were expected to find more attractive (smaller and with brighter blue eyes and more extensive red throat coloration). These females in turn had lower eigengene expression of the red module if they showed more interest later in courtship. Paxton benthic females showed the exact opposite relationships between red module eigengene expression and male morphology and female behaviour. The finding of overlap in modules showing associations with male traits and female behaviour suggests these modules are important for premating isolation.

Candidate gene expression also associated with trait values

Numerous candidate genes have previously been implicated as playing a role in female mate choice and social decision-making more generally [9,63]. We therefore mined our transcriptome dataset for candidate genes representing five distinct and well-studied neuroendocrine and neuromodulatory pathways. Specifically, we focused on genes involved in

1) **gonadotropin-releasing hormone (GnRH) signalling**, due to their role in reproduction [64]: *gnrh1, gnrh2, gnrh3, gnrhr4*; 2) **nonapeptide signalling**, known for regulating affiliative behaviour [65]: *avp, avpr2, oxt, oxtr*; 3) **dopamine signalling**, which is important for motivational processing [66]: *th, th2, DRD1, drd1b, drd2a, drd2l*; 4) **prostaglandin F2 alpha (PGF2α) signalling**, as the ovarian hormone PGF2α is a well-known regulator of reproductive behaviour in fishes [67,68]: *ptgfr*; and 5) specific genes important to **synaptic plasticity** that have previously been implicated in mate choice decisions in poecilid fishes [15,16,69]: *nlgn1, nlgn2a, nlgn2b, nlgn3a, nlgn3b, neuroligins, neuroserpin1*. We found that 20 of these 24 genes (83%) were members of 9 different gene co-expression modules (Fig. 7), with the blue (5 genes) and brown (4 genes) modules most prominently represented (recall that the brown module was correlated with Mbehav PC1, courtship vigour). We then used linear models as with the gene co-expression modules above to discover that 14 genes (58%), representing four of these pathways (all except for PGF2α signalling), showed significant differences in expression between populations (Fig. 7, Supplementary Table 9).

When we focused on significant relationships between candidate gene expression and specific male or female traits, we discovered several intriguing associations. For example, the expression of *gnrh3*, which has previously been shown to gate mating preferences in medaka (Japanese rice fish, *Oryzias latipes*) [70], was associated with female responsiveness (Fbehav PC1), while *gnrh2* expression reflected female timing (Fbehav PC2). Note that GnRH2 plays a critical role in the integration of energy homeostasis and sexual behaviour in mammals and teleosts [71,72]. Looking at the behaviour and morphology of the males, we found that

dopaminergic and nonapeptide signalling along with synaptic plasticity in the female brain reflect male courtship vigour (Mbehav PC1), while oxytocin receptor expression is related to male courtship strategy (Mbehav PC2). In addition, GnRH and dopaminergic signalling along with synaptic plasticity reflect a male's nuptial coloration (Mmorph PC1) and male size and body coloration (Mmorph PC2) with nonapeptide signalling also associated with Mmorph PC2. Taken together, our candidate gene analysis is consistent with findings in other species.

Conclusion

Despite their close evolutionary relationship, we find substantial differences in gene expression for limnetic and benthic females making mate choice decisions. This differentiation is not due simply to differences in magnitude of expression in the brains of the two species. Indeed, our most novel finding is that brain gene expression responds to male display traits in opposite directions for the two sympatric species, mirroring contrasting female behavioural responses to those displays that are known to contribute to sexual isolation. Male displays that trigger elevated expression of a module in female brains for one species trigger reduced expression in the other sympatric species. We find support not only of our expectation of constitutive expression differences in the brains of benthic and limnetic fish, but also our key prediction of differential expression driven by diverged female preferences for diverged male display traits in a quantitative manner related directly to female experience during courtship. Taken together, our results provide novel insights into the neuromolecular processes that govern reproductive isolation of diverging species.

Acknowledgements

Research was conducted under permits from the Ministry of the Environment, BC and approval of the Institutional Animal Care and Use Committee of Michigan State University. We thank J. Martinez, K. Doolittle, and B. Wurst for assistance on behavioural data collection and brain extractions, L. Racey, R. Mobley, J. Schuster, M. Tillotson, and E. Weigel for fish collection and/or fish care, and other members of the Boughman laboratory for additional support. A. Bell, H. Eisthen, J. Gallant, and M. Lucas gave valuable assistance while developing the brain extraction procedure. R. Harris extracted the RNA. D. Arasappan and B. Goetz gave bioinformatics training. B. Young gave bioinformatics training and extensive feedback on different versions of this manuscript. C. Weitecamp, C. Ofria, D. Bolnick, and M. Cummings gave other forms of support.

Funding

This research was supported with funding to JK, HAH, and JWB by the BEACON Center for the Study of Evolution in Action (NSF DBI-0939454) and to JWB by a CAREER grant from the National Science Foundation (DEB-0952659). JK was also supported by an Integrative Biology of Social Behavior Research Exchange Grant and the USDA National Institute of Food and Agriculture Federal Appropriations under project PEN04768 and accession no. 1026660.

References

1. Andersson M. 1994 *Sexual Selection*. Princeton University Press.

2. Panhuis TM, Butlin R, Zuk M, Tregenza T. 2001 Sexual selection and speciation. *Trends Ecol. Evol.* **16**, 364–371.

3. Mendelson TC, Safran RJ. 2021 Speciation by sexual selection: 20 years of progress. *Trends Ecol. Evol.* **36**, 1153–1163. (doi:10.1016/j.tree.2021.09.004)

4. Servedio MR, Boughman JW. 2017 The role of sexual selection in local adaptation and speciation. *Annu. Rev. Ecol. Evol. Syst.* **48**, 85–109. (doi:10.1146/annurev-ecolsys-110316-022905)

5. Jennions MD, Petrie M. 2007 Variation in mate choice and mating preferences: A review of causes and consequences. *Biol. Rev.* **72**, 283–327. (doi:10.1111/j.1469-185X.1997.tb00015.x)

6. Rosenthal GG, Ryan MJ. 2022 Sexual selection and the ascent of women: Mate choice research since Darwin. *Science* **375**, eabi6308. (doi:10.1126/science.abi6308)

7. Kraaijeveld K, Kraaijeveld-Smit FJL, Maan ME. 2011 Sexual selection and speciation: The comparative evidence revisited. *Biol. Rev.* **86**, 367–377. (doi:10.1111/j.1469-185X.2010.00150.x)

8. Ritchie MG. 2007 Sexual selection and speciation. *Annu. Rev. Ecol. Evol. Syst.* **38**, 79–102.

9. DeAngelis RS, Hofmann HA. 2020 Neural and molecular mechanisms underlying female mate choice decisions in vertebrates. *J. Exp. Biol.* **223**, jeb207324. (doi:10.1242/jeb.207324)

10. Ryan MJ. 2021 Darwin, sexual selection, and the brain. *Proc. Natl. Acad. Sci.* **118**, e2008194118. (doi:10.1073/pnas.2008194118)

11. Pfenning AR *et al.* 2014 Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science* **346**, 1256846. (doi:10.1126/science.1256846)

12. Renn SCP, Machado HE, Duftner N, Sessa AK, Harris RM, Hofmann HA. 2018 Gene expression signatures of mating system evolution. *Genome* **61**, 287–297. (doi:10.1139/gen-2017-0075)

13. Young RL *et al.* 2019 Conserved transcriptomic profiles underpin monogamy across vertebrates. *Proc. Natl. Acad. Sci.* **116**, 1331–1336. (doi:10.1073/pnas.1813775116)

14. Young RL, Weitekamp CA, Triki Z, Su Y, Bshary R, Hofmann HA. 2022 Shared neural transcriptomic patterns underlie the repeated evolution of mutualistic cleaning behavior in Labridae wrasses. *EcoEvoRxiv* **preprint**. (doi:https://doi.org/10.32942/osf.io/4p638)

15. Cummings ME, Larkins-Ford J, Reilly CRL, Wong RY, Ramsey M, Hofmann HA. 2008 Sexual and social stimuli elicit rapid and contrasting genomic responses. *Proc. R. Soc. B Biol. Sci.* **275**, 393–402. (doi:10.1098/rspb.2007.1454)

16. Bloch NI, Corral-López A, Buechel SD, Kotrschal A, Kolm N, Mank JE. 2018 Early neurogenomic response associated with variation in guppy female mate preference. *Nat. Ecol. Evol.* **2**, 1772–1781. (doi:10.1038/s41559-018-0682-4)

17. Harris RM, Hofmann HA. 2014 Neurogenomics of Behavioral Plasticity. In *Ecological Genomics* (eds CR Landry, N Aubin-Horth), pp. 149–168. Dordrecht: Springer Netherlands. (doi:10.1007/978-94-007-7347-9_8)

18. Morandin C *et al.* 2016 Comparative transcriptomics reveals the conserved building blocks involved in parallel evolution of diverse phenotypic traits in ants. *Genome Biol.* **17**, 43. (doi:10.1186/s13059-016-0902-7)

19. Taylor EB, McPhail J. 2000 Historical contingency and ecological determinism interact to prime speciation in sticklebacks, *Gasterosteus*. *Proc. R. Soc. Lond. B Biol. Sci.* **267**, 2375–2384. (doi:10.1098/rspb.2000.1294)

20. McKinnon JS, Mori S, Blackman BK, David L, Kingsley DM, Jamieson L, Chou J, Schluter D. 2004 Evidence for ecology's role in speciation. *Nature* **429**, 294–298. (doi:10.1038/nature02556)

21. McPhail J. 1994 Speciation and the evolution of reproductive isolation in the sticklebacks (*Gasterosteus*) of south-western British Columbia. In *The Evolutionary Biology of the Threespine Stickleback* (eds M Bell, S Foster), pp. 399–437. Oxford: Oxford Science Publications.

22. Rundle HD, Nagel L, Boughman JW, Schluter D. 2000 Natural selection and parallel speciation in sympatric sticklebacks. *Science* **287**, 306–308.

23. Boughman JW. 2001 Divergent sexual selection enhances reproductive isolation in sticklebacks. *Nature* **411**, 944–948. (doi:10.1038/35082064)

24. Keagy J, Lettieri L, Boughman JW. 2016 Male competition fitness landscapes predict both forward and reverse speciation. *Ecol. Lett.* **19**, 71–80. (doi:10.1111/ele.12544)

25. Nagel L, Schluter D. 1998 Body size, natural selection, and speciation in sticklebacks. *Evolution* **52**, 209–218. (doi:10.1111/j.1558-5646.1998.tb05154.x)

26. Kozak GM, Reisland M, Boughmann JW. 2009 Sex differences in mate recognition and conspecific preference in species with mutual mate choice. *Evolution* **63**, 353–365. (doi:10.1111/j.1558-5646.2008.00564.x)

27. Head ML, Kozak GM, Boughman JW. 2013 Female mate preferences for male body size and shape promote sexual isolation in threespine sticklebacks. *Ecol. Evol.* **3**, 2183–2196. (doi:10.1002/ece3.631)

28. Conte GL, Schluter D. 2013 Experimental confirmation that body size determines mate preference via phenotype matching in a stickleback species pair. *Evolution* , **67,** 1477-84. (doi:10.1111/evo.12041)

29. Mobley RB, Tillotson ML, Boughman JW. 2016 Olfactory perception of mates in ecologically divergent stickleback: population parallels and differences. *Evol. Ecol. Res.* **17**, 551–564.

30. Gow JL, Peichel CL, Taylor EB. 2006 Contrasting hybridization rates between sympatric three-spined sticklebacks highlight the fragility of reproductive barriers between evolutionarily young species. *Mol. Ecol.* **15**, 739–752. (doi:10.1111/j.1365-294X.2006.02825.x)

31. Bolnick DI, Barrett RDH, Oke KB, Rennison DJ, Stuart YE. 2018 (Non)parallel evolution. *Annu. Rev. Ecol. Evol. Syst.* **49**, 303–330. (doi:10.1146/annurev-ecolsys-110617-062240)

32. Magalhaes IS, Whiting JR, D'Agostino D, Hohenlohe PA, Mahmud M, Bell MA, Skúlason S, MacColl ADC. 2020 Intercontinental genomic parallelism in multiple three-spined stickleback adaptive radiations. *Nat. Ecol. Evol.* **5**, 251–261. (doi:10.1038/s41559-020-01341-8)

33. Jacobs A *et al.* 2020 Parallelism in eco-morphology and gene expression despite variable evolutionary and genomic backgrounds in a Holarctic fish. *PLOS Genet.* **16**, e1008658. (doi:10.1371/journal.pgen.1008658)

34. Tinghitella RM, Weigel EG, Head M, Boughman JW. 2013 Flexible mate choice when mates are rare and time is short. *Ecol. Evol.* **3**, 2820–2831. (doi:10.1002/ece3.666)

35. Love MI, Huber W, Anders S. 2014 Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550. (doi:10.1186/s13059-014-0550-8)

36. Langfelder P, Horvath S. 2008 WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559. (doi:10.1186/1471-2105-9-559)

37. Langfelder P, Horvath S. 2012 Fast R functions for robust correlations and hierarchical clustering. *J. Stat. Softw.* **46**, 1–17. (doi:10.18637/jss.v046.i11)

38. R Core Team. 2022 R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

39. Kolberg L, Raudvere U, Kuzmin I, Vilo J, Peterson H. 2020 gprofiler2 -- an R package for gene list functional enrichment analysis and namespace conversion toolset g:Profiler. *F1000Research* **9(ELIXIR)**, 709.

40. Storey JD, Tibshirani R. 2003 Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci.* **100**, 9440–9445. (doi:10.1073/pnas.1530509100)

41. Fox J, Weisberg S. 2019 *An R Companion to Applied Regression*. 3rd edn. Thousand Oaks CA: Sage. See https://socialsciences.mcmaster.ca/jfox/Books/Companion.

42. Lackey ACR, Boughman JW. 2014 Female discrimination against heterospecific mates does not depend on mating habitat. *Behav. Ecol.* **25**, 1256–1267. (doi:10.1093/beheco/aru111)

43. Härer A, Bolnick DI, Rennison DJ. 2021 The genomic signature of ecological divergence along the benthic-limnetic axis in allopatric and sympatric threespine stickleback. *Mol. Ecol.* **30**, 451–463. (doi:10.1111/mec.15746)

44. Conte GL, Arnegard ME, Best J, Chan YF, Jones FC, Kingsley DM, Schluter D, Peichel CL. 2015 Extent of QTL reuse during repeated phenotypic divergence of sympatric threespine stickleback. *Genetics* **201**, 1189–1200. (doi:10.1534/genetics.115.182550)

45. Jones FC *et al.* 2012 A genome-wide SNP genotyping array reveals patterns of global and repeated species-pair divergence in sticklebacks. *Curr. Biol.* **22**, 83–90. (doi:10.1016/j.cub.2011.11.045)

46. Taylor EB, McPhail JD. 1999 Evolutionary history of an adaptive radiation in species pairs of threespine sticklebacks (*Gasterosteus*): Insights from mitochondrial DNA. *Biol. J. Linn. Soc.* **66**, 271–291. (doi:10.1111/j.1095-8312.1999.tb01891.x)

47. Schluter D, McPhail JD. 1992 Ecological character displacement and speciation in sticklebacks. *Am. Nat.* **140**, 85–108. (doi:10.1086/285404)

48. Boughman JW, Rundle HD, Schluter D. 2005 Parallel evolution of sexual isolation in sticklebacks. *Evolution* **59**, 361–373.

49. Ryan MJ, Rosenthal GG. 2001 Variation and selection in swordtails. In *Model Systems in Behavioral Ecology* (ed L Dugatkin), pp. 133–148. Princeton: Princeton University Press.

50. Vamosi SM. 2002 Predation sharpens the adaptive peaks: Survival trade-offs in sympatric sticklebacks. *Ann. Zool. Fenn.* **39**, 237–248.

51. Mani GS, Clarke BC. 1997 Mutational order: A major stochastic process in evolution. *Proc. R. Soc. Lond. B Biol. Sci.* **240**, 29–37. (doi:10.1098/rspb.1990.0025)

52. Schluter D. 2009 Evidence for ecological speciation and Its alternative. *Science* **323**, 737–741. (doi:10.1126/science.1160006)

53. Mendelson TC, Martin MD, Flaxman SM. 2014 Mutation-order divergence by sexual selection: Diversification of sexual signals in similar environments as a first step in speciation. *Ecol. Lett.* **17**, 1053–1066.

54. Soria-Carrasco V *et al.* 2014 Stick insect genomes reveal natural selection's role in parallel speciation. *Science* **344**, 738–742. (doi:10.1126/science.1252136)

55. Endler JA, Houde AE. 1995 Geographic variation in female preferences and for male traits in *Poecilia reticulata*. *Evolution* **49**, 456–468. (doi:10.1111/j.1558-5646.1995.tb02278.x)

56. Scordato ES, Safran RJ. 2014 Geographic variation in sexual selection and implications for speciation in the Barn Swallow. *Avian Res.* **5**, 8. (doi:10.1186/s40657-014-0008-4)

57. Sinha S *et al.* 2020 Behavior-related gene regulatory networks: A new level of organization in the brain. *Proc. Natl. Acad. Sci.* **117**, 23270–23279. (doi:10.1073/pnas.1921625117)

58. Albert AYK, Millar NP, Schluter D. 2007 Character displacement of male nuptial colour in threespine sticklebacks (*Gasterosteus aculeatus*). *Biol. J. Linn. Soc.* **91**, 37–48. (doi:10.1111/j.1095-8312.2007.00777.x)

59. Lackey A, Boughman J. 2013 Divergent sexual selection via male competition: Ecology is key. *J. Evol. Biol.* **26**, 1611–1624.

60. Boughman JW. 2007 Condition-dependent expression of red colour differs between stickleback species. *J. Evol. Biol.* **20**, 1577–1590. (doi:10.1111/j.1420-9101.2007.01324.x)

61. Gerhardt HC, Brooks R. 2009 Experimental analysis of multivariate female choice in gray treefrogs (*Hyla versicolor*): Evidence for directional and stabilizing selection. *Evolution* **63**, 2504–2512. (doi:10.1111/j.1558-5646.2009.00746.x)

62. Oh KP, Shaw KL. 2013 Multivariate sexual selection in a rapidly evolving speciation phenotype. *Proc. R. Soc. B Biol. Sci.* **280**, 20130482. (doi:10.1098/rspb.2013.0482)

63. Weitekamp CA, Hofmann HA. 2017 Brain systems underlying social behavior. In *Evolution of Nervous Systems*, pp. 327–334. Elsevier. (doi:10.1016/B978-0-12-804042-3.00025-7)

64. Gore A. 2002 *GnRH: The Master Molecule of Reproduction*. Springer.

65. Goodson J. 2008 Nonapeptides and the evolutionary patterning of sociality. In *Progress in Brain Research*, pp. 3–15. Elsevier. (doi:10.1016/S0079-6123(08)00401-9)

66. Lerner TN, Holloway AL, Seiler JL. 2021 Dopamine, updated: Reward prediction error and beyond. *Curr. Opin. Neurobiol.* **67**, 123–130. (doi:10.1016/j.conb.2020.10.012)

67. Kidd MR, Dijkstra PD, Alcott C, Lavee D, Ma J, O'Connell LA, Hofmann HA. 2013 Prostaglandin F2α facilitates female mating behavior based on male performance. *Behav. Ecol. Sociobiol.* **67**, 1307–1315. (doi:10.1007/s00265-013-1559-9)

643 68. Juntti SA, Hilliard AT, Kent KR, Kumar A, Nguyen A, Jimenez MA, Loveland JL, Mourrain P,
644 Fernald RD. 2016 A neural basis for control of cichlid female reproductive behavior by
645 Prostaglandin F2α. *Curr. Biol.* **26**, 943–949. (doi:10.1016/j.cub.2016.01.067)

646 69. Wong RY, Ramsey ME, Cummings ME. 2012 Localizing brain regions associated with female
647 mate preference behavior in a swordtail. *PLoS ONE* **7**, e50355.
648 (doi:10.1371/journal.pone.0050355)

649 70. Okuyama T *et al.* 2014 A neural mechanism underlying mating preferences for familiar
650 individuals in medaka fish. *Science* **343**, 91–94. (doi:10.1126/science.1244724)

651 71. Temple JL, Millar RP, Rissman EF. 2003 An evolutionarily conserved form of gonadotropin-
652 releasing hormone coordinates energy and reproductive behavior. *Endocrinology* **144**, 13–
653 19. (doi:10.1210/en.2002-220883)

654 72. Marvel M, Levavi-Sivan B, Wong T-T, Zmora N, Zohar Y. 2021 Gnrh2 maintains reproduction
655 in fasting zebrafish through dynamic neuronal projection changes and regulation of
656 gonadotropin synthesis, oogenesis, and reproductive behaviors. *Sci. Rep.* **11**, 6657.
657 (doi:10.1038/s41598-021-86018-3)

658

659 **Figure Legends**

660 **Figure 1. Study Overview. (A) Experimental design.** Male and female stickleback fish were
661 sampled from two lakes on Texada Island, British Columbia and brought into the lab. Females
662 (right top corner of each panel) were exposed to three treatments before they were euthanized
663 and brains removed: heterospecific male courtship, conspecific male courtship, or a conspecific
664 female from their home aquarium. **(B) Female behaviour.** Females showed stronger preference
665 for conspecific males than heterospecific males, especially if the female was limnetic. Boxes
666 indicate means ± SE. Each circle is an individual female. Circles that are filled in indicate females
667 who were sampled for RNAseq. **(C) Male traits.** Five male morphological traits and five male
668 behavioural traits were collapsed into two principal components each to summarize male
669 variation. Male morphological PCs differentiated limnetic from benthic fish (Welch two-sample
670 t-test: Mmorph-PC1: $t_{52.0} = -4.13$, P = 0.00013, Mmorph-PC2: $t_{54.5} = -2.76$, P = 0.0078). The first
671 male behavioural PC (courtship vigour) did not distinguish species ($t_{52.06} = -1.39$, P = 0.17), but
672 the second one, which describes well known species differences in courtship strategy, did ($t_{54.73}$
673 = -3.49, P = 0.00095).

674

675 **Figure 2. Species have highly divergent gene expression patterns. (A) Principal component**
676 **analysis (PCA) of variance stabilized normalized counts of 90% most variable genes.** PC1
677 clearly separates populations in expected ways. For example, limnetic populations overlap and
678 are very well separated from the benthic population. The limnetic population that is most
679 similar to the benthic population is the one from the same lake. Ovals are 95% confidence
680 ellipses. Two samples appeared to have been swapped (Priest limnetic and Paxton benthic; they

are at the centre of the other's population distribution) and are not included in further analyses. **(B) Venn diagram comparing overlap of population comparison differentially expressed genes (DEGs, FDR corrected p-value < 0.1).** Again, there were many more genes that distinguished benthic from limnetic fish than limnetic fish from different lakes.

Figure 3. Differentially expressed genes from treatment comparisons. (A) UpSet plot showing all treatment comparisons separated by population. Solid circles connected by lines indicate the intersection between those sets – circles that are unconnected to any others indicate genes unique to that set. Generally, there is very little overlap between sets; when there is, it is between different comparisons within the same population or species. We used a significance threshold of FDR corrected $P < 0.1$. Traditional Venn Diagrams can be derived from the UpSet plot as seen in **(B) Paxton limnetic treatment comparisons** and **(C) Conspecific versus heterospecific comparisons.**

Figure 4. Comparisons of module expression between (A) populations and (B) treatments. Symbols and lines correspond to mean ± 95% confidence intervals for t-tests. Significant differences are indicated in colour.

Figure 5. Gene co-expression modules whose expression in female brains are predicted by male throat colour, but in different ways depending on population. Significance of the fixed effects of population, traits, and their interaction are indicated: *** < 0.001, ** < 0.01, * < 0.05.

Figure 6. Linear models describing relationship between module eigengene expression and population, traits, and their interaction. Arrows indicate modules that respond to both a male display trait and female courtship behaviour; these relationships are further explored in scatterplots. “P” = population term, “T” = trait term, “PxT” = population-by-trait interaction term with significance indicated: *** < 0.001, ** < 0.01, * < 0.05. To focus on the trait and population-by-trait interaction effects, we have reduced the population term to a single column indicating in how many of the six models it was significant (at $P < 0.05$).

Figure 7. Linear models describing relationship between candidate gene expression and population, traits, and their interaction. “P” = population term, “T” = trait term, “PxT” = population-by-trait interaction term, with significance indicated: *** < 0.001, ** < 0.01, * < 0.05. To focus on the trait and population-by-trait interaction effects, we have reduced the population term to a single column indicating in how many of the six models it was significant (at $P < 0.05$).